

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:08:07 ; Search time 36 Seconds

(without alignments)  
59.222 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYASALGSDGVNVT 16

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SID52/gcgdata/geneseq/genesep-emb1/AA1980.DAT.\*  
2: /SID52/gcgdata/geneseq/genesep-emb1/AA1981.DAT.\*  
3: /SID52/gcgdata/geneseq/genesep-emb1/AA1982.DAT.\*  
4: /SID52/gcgdata/geneseq/genesep-emb1/AA1983.DAT.\*  
5: /SID52/gcgdata/geneseq/genesep-emb1/AA1984.DAT.\*  
6: /SID52/gcgdata/geneseq/genesep-emb1/AA1985.DAT.\*  
7: /SID52/gcgdata/geneseq/genesep-emb1/AA1986.DAT.\*  
8: /SID52/gcgdata/geneseq/genesep-emb1/AA1987.DAT.\*  
9: /SID52/gcgdata/geneseq/genesep-emb1/AA1988.DAT.\*  
10: /SID52/gcgdata/geneseq/genesep-emb1/AA1989.DAT.\*  
11: /SID52/gcgdata/geneseq/genesep-emb1/AA1990.DAT.\*  
12: /SID52/gcgdata/geneseq/genesep-emb1/AA1991.DAT.\*  
13: /SID52/gcgdata/geneseq/genesep-emb1/AA1992.DAT.\*  
14: /SID52/gcgdata/geneseq/genesep-emb1/AA1993.DAT.\*  
15: /SID52/gcgdata/geneseq/genesep-emb1/AA1994.DAT.\*  
16: /SID52/gcgdata/geneseq/genesep-emb1/AA1995.DAT.\*  
17: /SID52/gcgdata/geneseq/genesep-emb1/AA1996.DAT.\*  
18: /SID52/gcgdata/geneseq/genesep-emb1/AA1997.DAT.\*  
19: /SID52/gcgdata/geneseq/genesep-emb1/AA1998.DAT.\*  
20: /SID52/gcgdata/geneseq/genesep-emb1/AA1999.DAT.\*  
21: /SID52/gcgdata/geneseq/genesep-emb1/AA2000.DAT.\*  
22: /SID52/gcgdata/geneseq/genesep-emb1/AA2001.DAT.\*  
23: /SID52/gcgdata/geneseq/genesep-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	20	AAU25521
2	49	62.8	262	22	AAU34559
3	49	62.8	262	22	AAU02203
4	49	62.8	262	22	AAU38231
5	49	62.8	429	20	AAU25520
6	49	62.8	491	22	ABG28562
7	49	62.8	693	22	ABG28564
8	48	61.5	267	22	AAU08027
9	48	61.5	537	22	AAU17777
10	48	61.5	1050	22	ABG10171

11	48	61.5	1053	22	AAU39991
12	45	57.7	357	21	AAU16520
13	44	56.4	265	22	AAU36258
14	44	56.4	268	18	AAU99839
15	41	52.6	415	23	ABU93255
16	41	52.6	526	10	AAU90387
17	40	51.3	33	22	AAU11660
18	40	51.3	46	22	AAU11393
19	40	51.3	51	23	ABU65185
20	40	51.3	402	23	AAU76528
21	40	51.3	425	22	AAU09883
22	40	51.3	425	22	ABU52493
23	39	50.0	64	22	AAU08242
24	39	50.0	70	22	AAU27464
25	39	50.0	183	21	AAU08315
26	39	50.0	183	21	AAU58516
27	39	50.0	224	20	AAU97727
28	39	50.0	261	21	AAU03632
29	39	50.0	261	21	AAU03635
30	39	50.0	261	21	AAU03635
31	39	50.0	262	23	ABU48921
32	39	50.0	279	21	AAU40867
33	39	50.0	315	21	AAU40866
34	39	50.0	325	21	AAU40865
35	39	50.0	2392	21	AAU07665
36	38	48.7	41	22	AAU13380
37	38	48.7	46	22	AAU09289
38	38	48.7	88	22	AAU09624
39	38	48.7	123	21	AAU47869
40	38	48.7	144	21	AAU47868
41	38	48.7	157	19	AAU62720
42	38	48.7	170	21	AAU47867
43	38	48.7	218	21	AAU14357
44	38	48.7	219	21	AAU47843
45	38	48.7	239	21	AAU14356

#### ALIGNMENTS

RESULT 1	AAU25521	standard; Protein: 262 AA.
XX	AAU25521	
AC	AAU25521	
XX		
DT	30-SEP-1999	(first entry)
XX		
DE	E. coli mutant FabI protein.	
XX		
KW	FabI; enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy; fatty acid biosynthesis; detection; antibacterial; antifungal; treatment; antiprotocoll; antiparasitic; antiviral; soap; infection; disinfectant;	
KW	dergent; mouthwash; toothpaste; contraceptive; inhibitor; trichosan;	
KW	resistance; NSAM; non-specific antimicrobial; mutant.	
XX		
OS	Escherichia coli.	
XX		
XX		
FT	Key	Location/Qualifiers
FT	Protein	1..262
FT		/label="FabI
FT		/note="No stop codon given in specification"
XX		
PN	W09937800-A1.	
XX		
PD	29-JUL-1999.	
XX		
PF	22-JAN-1999;	99WO-US01288.
XX		
PR	26-JAN-1998;	98US-0013440.
XX	23-JAN-1998;	98US-0072244.
XX		

PA (TUFT ) TUFTS COLLEGE.  
 XX  
 PI Levy SB, McMurry LM;  
 XX  
 DR WPI: 1999-458702/38.  
 DR N-PSDB: AAX88348.  
 XX  
 PT Identifying antimicrobials that target enoyl acyl carrier protein  
 PT reductase, for therapeutic use and for incorporation into e.g. soaps  
 PT and detergents  
 XX  
 PS Claim 77, Fig 2; 80pp; English.  
 XX  
 CC This invention describes a novel method for identifying antimicrobials  
 CC (1) that interact with an enoyl acyl carrier protein (ACP) reductase  
 CC (ER) polypeptide, which is essential for fatty acid biosynthesis, by  
 CC contacting ER with test compound and detecting any interaction. The  
 CC identified antimicrobials have antibacterial, antifungal, antiprotocozal,  
 CC antiparasitic and antiviral activity. The products of the invention are  
 CC used (1) therapeutically to treat a wide variety of viral, bacterial,  
 CC fungal, yeast and protozoal infections, in both humans and animals, and  
 CC (11) in disinfectants, detergents, soaps, mouthwashes, toothpastes and  
 CC also contraceptive devices. The invention describes methods for (1)  
 CC identifying a (1) from its effect on enzymatic activity of ER, (2)  
 CC identifying a (1) from its ability to inhibit fatty acid biosynthesis in  
 CC a microbe, (3) identifying a (1) that interacts with a mutant ER, (4)  
 CC identifying (1) that inhibit proliferation or viability of a microbe that  
 CC is resistant to triclosan or to NSAM (non-specific antimicrobials). This  
 CC sequence represents a mutant Escherichia coli FabI protein which  
 CC interacts with the enoyl-acyl carrier protein (ACP) reductase to  
 CC illustrate the method of the invention.  
 XX  
 SQ Sequence 262 AA;  
 Query Match 62.8%; Score 49; DB 20; Length 262;  
 Best Local Similarity 60.0%; Pred. No. 2.2;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 MRYRASALGSDGVRY 15  
 DB 170 VRYMANMGPGEVRY 184  
 RESULT 2  
 AAU34559  
 ID AU34559 standard; Protein: 262 AA.  
 XX  
 AC AU34559;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE E. coli cellular proliferation protein #140.  
 XX  
 KW Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX  
 OS Escherichia coli.  
 XX  
 FN WO200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 XX 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.

XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI: 2001-611495/70.  
 DR N-PSDB: AAS52418.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID No 10152; 511pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 262 AA;  
 Query Match 62.8%; Score 49; DB 22; Length 262;  
 Best Local Similarity 60.0%; Pred. No. 2.2;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 MRYRASALGSDGVRY 15  
 DB 170 VRYMANMGPGEVRY 184  
 RESULT 3  
 AAE02203  
 ID AAE02203 standard; Protein: 262 AA.  
 XX  
 AC AAE02203;  
 XX  
 DT 31-JUL-2001 (first entry)  
 XX  
 DE Escherichia coli NADPH-dependent enoyl-ACP reductase (FabI).  
 XX  
 KW FabI, high throughput method; fatty acid biosynthesis; therapy;  
 KW bacterial enzyme; biological agent screening; otitis media; empyema;  
 KW bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;  
 KW infective endocarditis; secretory diarrhoea; splenic; retroperitoneal;  
 KW intrarenal; perinephric; cutaneous; abscess; blepharitis;  
 KW conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;  
 KW cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndrome;  
 KW impetigo; folliculitis; wound infection; bacterial myositis;  
 KW septic arthritis; osteomyelitis; enoyl-ACP reductase;  
 KW acyl carrier protein.  
 XX  
 OS Escherichia coli.  
 XX  
 PD WO200130988-A1.  
 XX  
 PD 03-MAY-2001.  
 XX  
 PF 26-OCT-2000; 2000WO-US29451.  
 XX  
 PR 27-OCT-1999; 99US-0161775.

XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Dewolf W, Kallender H, Lonsdale JR;  
 XX WPI: 2001-316332/33.  
 DR N-PSDB: AAD06214.  
 XX  
 PT High throughput method for screening for biological agents against  
 PT fatty acid biosynthesis comprises contacting a bacterial enzymatic  
 PT pathway with enzymes e.g. malonyl-CoA:ACP transacylase -  
 XX  
 PS Claim 1: Page 15; 94pp; English.  
 XX  
 CC The present invention relates to a high throughput method for screening  
 CC biological agents affecting fatty acid biosynthesis, comprises  
 CC contacting a bacterial enzymatic pathway with enzymes. The method is  
 CC used for screening biological agents affecting fatty acid biosynthesis.  
 CC Agonists and antagonists of fab (fatty acid biosynthesis) are used to  
 CC inhibit, prevent or treat diseases such as infections of the upper  
 CC respiratory tract (e.g. otitis media, bacterial tracheitis, acute  
 CC epiglottitis, thyroiditis), lower respiratory (e.g. empyema, lung  
 CC abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g.  
 CC secretory diarrhoea, splenic abscess, retroperitoneal abscess), central  
 CC nervous system (e.g. cerebral abscess), eye (e.g. blepharitis,  
 CC conjunctivitis, keratitis, endophthalmitis, preseptal and orbital  
 CC cellulitis, dacryocystitis), kidney and urinary tract (e.g.  
 CC epididymitis, intrarenal and perinephric abscess, toxic shock syndrome),  
 CC skin (e.g. impetigo, folliculitis, cutaneous abscess, cellulitis, wound  
 CC infection, bacterial myositis) and bone and joint (e.g. septic  
 CC arthritis, osteomyelitis).  
 CC The present sequence is Escherichia coli MADPH-dependent  
 CC enoyl-ACP (acyl carrier protein) reductase (FabI). In fatty acid  
 CC biosynthetic pathway, dehydration by FabZ leads to trans-2-enoyl-ACP  
 CC which is in turn converted to acyl-ACP by FabI.  
 XX  
 SQ Sequence 262 AA:  
 Query Match 62.8%; Score 49; DB 22; Length 262;  
 Best Local Similarity 60.0%; Pred. No. 2.2;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 MRYRASALGSDGVRV 15  
 :|||:|:|:|:|  
 Db 170 VRYMANMGPCEVRV 184

RESULT 4  
 AAU38231  
 ID AAU38231 standard; Protein: 269 AA.  
 AC AAU38231;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Salmonella typhi cellular proliferation protein #122.  
 XX  
 KW Antisense: prokaryotic cellular proliferation protein;  
 KW antibiotic: antibacterial; drug design.  
 XX  
 OS Salmonella typhi.  
 XX  
 PM MO200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207227P.  
 PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELITR) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawlick JD, Carr CJ;  
 PI Yamamoto RT, Xu HH;  
 XX WPI: 2001-611495/70.  
 DR N-PSDB: AAS56090.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID No 13824; 511pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 269 AA:  
 Query Match 62.8%; Score 49; DB 22; Length 269;  
 Best Local Similarity 60.0%; Pred. No. 2.3;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 MRYRASALGSDGVRV 15  
 :|||:|:|:|:|  
 Db 177 VRYMANMGPCEVRV 191

RESULT 5  
 AAY25520  
 ID AAY25520 standard; Protein: 429 AA.  
 AC AAY25520;  
 XX  
 DT 30-SEP-1999 (first entry)  
 XX  
 DE E. coli FabI protein fragment.  
 XX  
 KW FabI: enoyl ACP reductase; acyl carrier protein; antimicrobial; therapy;  
 KW fatty acid biosynthesis; detection; antibacterial; antifungal; treatment;  
 KW antiprotozoal; antiparasitic; antiviral; soap; infection; disinfectant;  
 KW detergent; mouthwash; toothpaste; contraceptive; inhibitor; triclosan;  
 KW resistance; NSAM; non-specific antimicrobial.  
 XX  
 OS Escherichia coli.  
 XX  
 FH Key  
 FH Protein 1..429  
 FT Location/Qualifiers  
 FT /label= FabI  
 FT /note= "No start or stop codon given in specification"  
 FT Misc-difference 74  
 FT /note= "in-frame stop codon encoded by TGA"  
 FT Misc-difference 111



PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
PI Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS92751.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20; SEQ ID NO 58923; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 693 AA:  
Query Match 62.8%; Score 49; DB 22; Length 693;  
Best Local Similarity 60.0%; Pred. No. 6.2;  
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
OY 1 MRYRASALGSDGVRV 15  
:|||:|:|:|:|:|:|  
Db 601 VRYMANAMGPEGVRY 615  
RESULT 8  
AAB90827  
ID AAB90827 standard; Protein: 267 AA.  
XX  
AC AAB90827;  
XX  
DT 15-JUN-2001 (first entry)  
XX  
DE Human shear stress-response protein SEQ ID NO: 171.  
XX  
KW Human; shear stress-response protein; vascular disease;  
KW arteriosclerosis.  
XX  
OS Homo sapiens.  
XX  
MO200125427-A1.  
XX  
PN 12-APR-2001.  
XX  
PD  
XX

PF 02-OCT-2000; 2000WO-JP06840.  
XX  
XX 01-OCT-1999; 99JP-0280976.  
PR  
XX (KYOM) KYOMA HAKKO KOGYO KK.  
PA (NOJI/) NOJIMA H.  
PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;  
PI Kuga T, Sekine S, Nakamura Y, Sugano S;  
XX WPI; 2001-266308/72.  
DR N-PSDB; AAH02967.  
XX  
PT DNA sequences, proteins encoded by them and antibodies against them  
PT useful in diagnosis and treatment of vascular disease caused by  
PT arteriosclerosis -  
XX  
XX  
PS Claim 35; Page 664-665; 678pp; Japanese.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human shear stress response proteins. These are useful in the  
CC diagnosis, treatment and screening of vascular diseases caused by  
CC arteriosclerosis, including heart failure, post-PTCA restenosis and  
CC hypertension.  
XX  
SQ Sequence 267 AA:  
Query Match 61.5%; Score 48; DB 22; Length 267;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 7 ALGSDGVRVT 16  
|||||||  
Db 26 ALGSDGVRVT 35  
RESULT 9  
AAM41777  
ID AAM41777 standard; Protein: 537 AA.  
XX  
AC AAM41777;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6708.  
XX  
KW Human; nocturnal immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CMS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
MO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX

PI	Mang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J:
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI; 2001-442253/47.
DR	N-PSDB; AAI60933.
XX	
PT	Novel nucleic acids and polypeptides, useful for treating disorders
PT	such as central nervous system injuries -
PS	Example 2; SEQ ID NO 6708; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilistation of the activities such as: immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
XX	specification.
SQ	
SQ	Sequence 537 AA:
Query Match	61.5%; Score 48; DB 22; Length 537;
Best Local Similarity	100.0%; Pred. No. 7;
Matches 10; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	7 ALGSDGVRYT 16 
Db	60 ALGSDGVRYT 69
RESULT 10	
ABGI0171	
ID	ABGI0171 standard; Protein; 1050 AA.
XX	
AC	ABGI0171;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #10162.
XX	
KM	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; diagnostic disorder.
XX	
OS	Homo sapiens.
XX	
WO	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
Pf	30-MAR-2001; 2001WO-US08631.
XX	
PR	31-MAR-2000; 2000US-0540217.
XX	
PA	23-AUG-2000; 2000US-0649167.
XX	
(HYSE-)	HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
XX	
DR	N-PSDB; AAS74358.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess

XX biodiversity -  
XX  
PS Claim 20: SEQ ID No 40530; 1033p; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 1050 AA.

Query Match	61.5%	Score 48;	DB 22;	length 1050;
Best Local Similarity	100.0%;	Pred. No. 14;		
Matches 10;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	7 ALGSDGVKRV 16			
DB	12 ALGSDGVKRV 21			
RESULT 11				
AAM39991				
ID	AAM39991 standard; Protein; 1063 AA.			
XX				
XX				
AC	AAM39991;			
XX				
DT	22-OCT-2001 (first entry)			
XX				
DE	Human polypeptide SEQ ID NO 3136.			
XX				
KW	Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;			
KW	peripheral nervous system; neuropathy; central nervous system; CNS;			
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;			
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemolactic;			
KW	chemokine; thrombolytic; drug screening; arthritis; inflammation;			
KW	leukemia.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200153312-A1.			
XX				
PD	26-JUL-2001.			
XX				
PF	26-DEC-2000; 2000WO-US34263.			
XX				
PR	21-JAN-2000; 2000US-0488725.			
PR	25-APR-2000; 2000US-0552317.			
PR	09-JUL-2000; 2000US-0598042.			
PR	19-JUL-2000; 2000US-0620312.			
PR	03-AUG-2000; 2000US-0653450.			
PR	14-SEP-2000; 2000US-0662191.			
PR	19-OCT-2000; 2000US-0693036.			
PR	29-NOV-2000; 2000US-0727344.			
XX				
XX				
PA	(HYSE-) HYSEQ INC.			
XX				

PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR MPI: 2001-442253/47.  
DR N-PSDB: AA159147.  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Example 4: SEQ ID NO 3136; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA138642-AA142213) with nootropic,  
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SO Sequence 1063 AA:  
Query Match 61.5%; Score 48; DB 22; Length 1063;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 7 AUGSCGCVRYT 16  
- | | | | | | | | | |  
Db 26 AUGSCGCVRYT 35  
RESULT 12  
AAG16520  
ID AAG16520 standard; Protein: 357 AA.  
XX  
XX AAG16520;  
AC  
XX  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17197.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PN  
XX  
XX 06-SEP-2000.  
PD  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123160.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137232.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145813.  
 PR 27-JUL-1999; 99US-0145818.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148365.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154039.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 57.7%; Score 45; DB 21; Length 357;  
 Best Local Similarity 69.2%; Pred. No. 15;  
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 RASALGDCVRYV 16  
 ||||| : : : : :  
 DB 43 RASALGNTGLKVT 55

## RESULT 13

AAU36258  
 ID AAU36258 standard; Protein; 265 AA.

AC AAU36258;  
 XX

DT 14-FEB-2002 (first entry)  
 XX

DE Pseudomonas aeruginosa cellular proliferation protein #248.  
 XX

KW Antisense; prokaryotic cellular proliferation protein;  
 XX antibiotic; antibacterial; drug design.

OS Pseudomonas aeruginosa.  
 XX

PN WO200170955-A2.  
 XX

PD 27-SEP-2001.  
 XX

PF 21-MAR-2001; 2001WO-US09180.  
 XX

PR 21-MAR-2000; 2000US-191078P.  
 XX

PR 23-MAY-2000; 2000US-206848P.  
 XX

PR 26-MAY-2000; 2000US-207727P.  
 XX

PR 23-OCT-2000; 2000US-242578P.  
 XX

PR 27-NOV-2000; 2000US-253625P.  
 XX

PR 22-DEC-2000; 2000US-257931P.  
 XX

PR 16-FEB-2001; 2001US-269308P.  
 XX

PA (ELIT- ) ELITRA PHARM INC.  
 XX

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS54117.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX



PS Example 3: Seq ID NO 11851: 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*

CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The

CC invention is also useful for the identification of potential new targets

CC for antibiotic development. The antisense nucleic acids can also be used

CC and to identify proteins used in proliferation, to express these proteins,

CC and to obtain antibodies capable of binding to the expressed proteins.

CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in

CC a wide variety of organisms. The present sequence represents an

CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences.

XX

SQ Sequence 265 AA:

Query Match 56.4%; Score 44; DB 22; Length 265;

Best Local Similarity 53.3%; Pred. No. 16;

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15

DB 173 VRYLASLAGEGTRV 187

||| | :||| :|||

RESULT 14

ID AAM9839

AAW9839 standard; Protein: 268 AA.

XX

AC- AAW9839;

XX

DT 08-JUN-1999 (first entry)

XX

DE L. pneumophila FabI enoyl reductase homologue protein sequence.

XX

KW Legionella pneumophila; FabI; enoyl reductase; antimicrobial;

KM gemfibrozil; growth inhibition; bacterium; infection.

XX

OS Legionella pneumophila.

XX

PN WO9731530-A1.

XX

PD 04-SEP-1997.

XX

PF 28-FEB-1997; 97WO-US03158.

XX

PR 29-FEB-1996; 96US-0608712.

XX

PA (UYCO ) UNIV COLUMBIA NEW YORK.

XX

PI Della-Latta P, Kabbash C, Shuman HA, Silverstein SC;

XX

DR WPI: 1997-448377/41.

XX

DR N-PSDB: AXI9775.

XX

PT Inhibiting growth of bacteria - by contacting the bacteria with

PT gemfibrozil or related compound

XX

PS Example 2: Fig 21B: 109pp; English.

XX

CC A method has been developed for: (i) inhibiting growth of a bacterium;

CC (ii) alleviating the symptoms of a bacterial infection in a subject;

CC (iii) inhibiting activity of enoyl reductase enzyme in a cell; or (iv)

CC altering a pathway of fatty acid synthesis in a bacterium, comprises

CC contacting the bacterium, subject or cell with a novel compound (I) or

CC its salt or ester. The above processes are especially applicable to

CC bacteria/bacterial cells selected from *Legionella pneumophila*,

CC *Mycobacterium tuberculosis*, *Bacillus subtilis*, *Bacillus megaterium*,

CC *Pseudomonas oleovorans*, *Alcaligenes eutrophus*, *Rhodococcus sp.*,

CC *Citrobacter freundii*, Group A *Streptococcus sp.*, *Coxsack* *Staphylococcus*

CC *aureus* or *Neisseria sp.* The bacterial infection is e.g. leprosy, brucella

CC or *Salmonella*. The present sequence represents a FabI enoyl reductase

CC homologue from *L. pneumophila*, which is used in an example from the

CC present invention.

XX

SQ Sequence 268 AA:

Query Match 56.4%; Score 44; DB 18; Length 268;

Best Local Similarity 53.3%; Pred. No. 16;

Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15

DB 177 VRYLASLGRGLKI 191

||| | :||| :|||

RESULT 15

ID ABB93255

ABB93255 standard; Protein: 415 AA.

XX

AC ABB93255;

XX

DT 31-MAY-2002 (first entry)

XX

DE Herbicidally active polypeptide SEQ ID NO 2466.

XX

KW Herbicidal; plant; agriculture; herbicide.

XX

OS Arabidopsis thaliana.

XX

PN WO200210210-A2.

XX

PD 07-FEB-2002.

XX

PF 28-AUG-2001; 2001WO-EP09892.

XX

PR 28-AUG-2001; 2001WO-EP09892.

XX

PA (FARB ) BAYER AG.

XX

PI Tietjen K, Weidner M;

XX

DR WPI: 2002-269010/31.

XX

PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences

PT from plant with nucleic acid or amino acid sequences from non-plant

PT organisms -

XX

PS Claim 5; SEQ ID NO 2466; 261pp + Sequence Listing; English.

XX

XX

CC The invention relates to identifying target proteins

CC (ABB90790-ABB94016) for herbicidally active compounds, comprising

CC aligning and comparing nucleic acid or amino acid sequences from plant

CC with nucleic acid or amino acid sequences from non-plant organisms using

CC suitable search parameters, where plant sequences having an E-value

CC greater by a factor of 3 than the E-value of most similar non-plant

CC sequences are selected. The polypeptides or nucleic acids encoding them

CC are useful for identifying modulators. The identified modulators are

CC useful as herbicides.

XX

SQ Sequence 415 AA:

Query Match 52.6%; Score 41; DB 23; Length 415;

Best Local Similarity 53.3%; Pred. No. 82;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 16

Thu Jan 16 15:49:08 2003

us-09-893-371-2.rag

Page 10

||:||||:|  
Db 392 RYKAGALGAERKRAT 406

Search completed: January 16, 2003, 10:12:11  
Job time : 38 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:10:53 ; Search time 29 seconds  
(without alignments)  
113.681 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDGVAVT 16

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_minc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeophages:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	100.0	1044	11	Q9ERB6 mus musculus
2	50	64.1	260	16	Q9CP73 pasteurella
3	49	62.8	262	16	Q87C7 salmonella
4	48	61.5	99	2	Q51891 proteus mir
5	45	57.7	264	16	Q8Y072 raietonia s
6	45	57.7	444	5	Q8WQ34 leishmania
7	44	55.4	268	2	Q9RE03 legionella
8	44	55.1	826	2	Q9EIK5 burkholderi
9	42	53.8	291	10	Q9LPE0 arabidopsis
10	41	52.6	151	2	Q9L5F6 salmonella
11	41	52.6	151	16	Q935M8 salmonella
12	41	52.6	252	16	Q92M00 rhizobium m
13	41	52.6	258	16	Q930L5 rhizobium m
14	41	52.6	356	10	Q81792 arabidopsis
15	41	52.6	415	10	Q9M068 arabidopsis
16	41	52.6	419	10	Q8V207 arabidopsis

17	41	52.6	572	16	Q9S1S8 streptomyc
18	40	51.3	375	16	Q9F2K4 streptomyc
19	40	51.3	412	4	Q96PW2 homo sapien
20	40	51.3	437	16	Q9KY22 streptomyc
21	40	51.3	687	5	Q9GTX9 armigeres s
22	39	50.0	242	2	Q9L8E7 vibrio harv
23	39	50.0	261	16	Q9K151 neisseria m
24	39	50.0	261	16	Q9JSS8 neisseria m
25	39	50.0	262	16	Q92D51 listeria in
26	39	50.0	262	16	Q918D5 listeria in
27	39	50.0	274	16	Q8YQD6 anabaena sp
28	39	50.0	467	10	Q8S1Q8 oryza sativ
29	39	50.0	551	2	Q8RR23 streptomyc
30	39	50.0	793	2	Q95485 lactobacill
31	39	50.0	793	2	Q48571 lactobacill
32	39	50.0	844	10	Q94BN8 arabidopsis
33	39	50.0	875	10	Q9FLQ1 arabidopsis
34	39	50.0	924	16	P73172 synchocyst
35	39	50.0	1296	16	Q9P9W1 xyella fas
36	39	50.0	1470	5	Q9G993 leishmania
37	39	50.0	2393	2	Q925F4 myxococcus
38	38	49.4	513	16	Q8X519 escherichia
39	38	48.7	110	16	Q92SV3 rhizobium m
40	38	48.7	146	10	Q64874 arabidopsis
41	38	48.7	192	5	Q8T692 giardia lam
42	38	48.7	254	16	Q92PP8 rhizobium m
43	38	48.7	272	16	Q8UHC5 agrobacteri
44	38	48.7	298	10	Q9FJ82 arabidopsis
45	38	48.7	312	16	Q97Q18 streptococc

## ALIGNMENTS

### RESULT 1

Q9ERB6 PRELIMINARY; PRT; 1044 AA.

AC Q9ERB6; 01-MAR-2001 (TREMURel. 16, Created)

DT 01-MAR-2001 (TREMURel. 16, Last sequence update)

DT 01-JUN-2002 (TREMURel. 21, Last annotation update)

DE Nuclear myosin I beta.

GN MYOIC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Pestle-Dragovich L., Stojiljkovic L., Phillimonenko A.A., Nowak G.,

RA Ke Y., Settlage R.E., Shabanowitz J., Hunt D.F., Hozak P.,

RA de Lanerolle P.,

RT "A Myosin I Isoform in the Nucleus."

RL Science 0:0-0(2000).

DR EMBL; AY007255; AAG02570.1; -

DR HSSP; P08799; 1MND.

DR MGD; MGI:106612; MYOIC.

DR InterPro; IPR000048; IQ\_region.

DR InterPro; IPR001609; myosin\_head.

DR Pfam; PF00612; IQ\_3.

DR Pfam; PF00663; myosin\_head; 1.

DR PRINTS; PR00193; MYOSINHEAVY.

DR ProDom; PD000355; myosin\_head; 1.

DR SMART; SM00015; IQ\_2.

DR SMART; SM00442; MISC; 1.

SQ SEQUENCE 1044 AA; 119876 MW; CDAB0FF699D041C9 CRC64;

Query Match 100.0%; Score 78; DB 11; Length 1044;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 16; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVAVT 16  
|||||

Db 1 MRYRASALGSDGVRV 16

# RESULT 2

09CP73

ID 09CP73 PRELIMINARY; PRT; 260 AA.

AC 09CP73: (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Pasteurella multocida.

GN FABI OR PM0182.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Pasteurella.

OX NCBI\_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PM70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RT "Complete genome sequence of Pasteurella multocida PM70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

DR EMBL; AF006052; AK02266.1; -.

DR HSSP; P29132; 10SG.

DR InterPro; IPR002198; ADH\_short.

DR Pfam; PF00106; adh\_short; 1.

KW Oxidoreductase; Complete proteome.

SO SEQUENCE 260 AA; 27850 MW; 922E5EBE8B0C8963 CRC64;

Query Match 64.1%; Score 50; DB 16; Length 260;

Best Local Similarity 64.3%; Pred. No. 1.7;

Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVRV 15

Db 171 RFMAALGKDGIRV 184

1:|||||1111

# RESULT 3

0827C7

ID 0827C7 PRELIMINARY; PRT; 262 AA.

AC 0827C7:

DT 01-MAR-2002 (TREMblrel. 20, Created)

DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Enoyl-[acyl-carrier-protein] reductase (NADH).

GN Stryl352.

OS Salmonella typhi.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

OX NCBI\_TaxID=601;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CT18;

RX MEDLINE=21534947; PubMed=11677608;

RA Pakhilli J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,

RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Feltwell T., Hamlin N., Hague A., Hen T.T., Holtroyd S., Jagels K.,

RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gea P., Parry C.,

RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,

RA Whitehead S., Barrett B.G.;

RT "Complete genome sequence of a multiple drug resistant Salmonella

enterica serovar Typhi CT18.";

Nature 413:848-852(2001).

EMBL; AL672720; CAP01621.1; -.

DR InterPro; IPR002198; ADH\_short.

DR Pfam; PF00106; adh\_short; 1.

KW Complete proteome.

SO SEQUENCE 262 AA; 27775 MW; 1844E8792E709A93 CRC64;

Query Match 62.8%; Score 49; DB 16; Length 262;

Best Local Similarity 60.0%; Pred. No. 2.6;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRV 15

Db 170 VRYMANMGPEGIRV 184

|||||1111

# RESULT 4

051891

ID 051891 PRELIMINARY; PRT; 99 AA.

AC 051891:

DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Similar to E. coli Envm (Fragment).

OS Proteus mirabilis.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Proteus.

OX NCBI\_TaxID=584;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AF 2924;

RX MEDLINE=93170270; PubMed=8436105;

RA Mignogna G., Allocali N., Aceto A., Piccolomini R., Di Ilio C.,

RA Baria D., Martini F.;

RT "The amino acid sequence of glutathione transferase from Proteus

mirabilis, a prototype of a new class of enzymes.";

RL Fur. J. Biochem. 211:421-425(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AF 2924;

RX MEDLINE=96358500; PubMed=8761466;

RA Perito B., Allocali N., Aceto A., Casalone E., Di Ilio C., Masulli M.,

RA Dragani B., Polinelli M.;

RT "Molecular cloning and overexpression of a glutathione transferase

RT gene from Proteus mirabilis.";

RL Biochem. J. 318:157-162(1996).

DR EMBL; U38482; AAC44361.1; -.

DR HSSP; P29132; 1DFT.

DR InterPro; IPR002198; ADH\_short.

DR Pfam; PF00106; adh\_short; 1.

FT NON\_TER 1

SO SEQUENCE 99 AA; 10395 MW; 2E442ABB04FA031 CRC64;

QY 1 MRYRASALGSDGVRV 15

Db 7 VRYMANMGPEGIRV 21

|||||1111

# RESULT 5

08Y072

ID 08Y072 PRELIMINARY; PRT; 264 AA.

AC 08Y072:

DT 01-MAR-2002 (TREMblrel. 20, Created)

DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Probable enoyl-[acyl-carrier-protein] reductase (EC 1.3.1.9).

GN FABI OR RSC1172 OR RS04528.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI\_TaxID=305;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GM11000;



DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE F6N18.2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 RN NCBI\_TaxID=3702;  
 RP SEQUENCE FROM N.A.  
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
 RA Shun P., Altai H., Bei Q., Chin C., Chou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaverl A.,  
 RA Tottum M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F6N18 from chromosome  
 I.";  
 RL EMBL: AC017118; AAF25981.1; -;  
 DR EMBL: AC017118; AAF25981.1; -;  
 SO SEQUENCE 291 AA; 32361 MW; 3F64694AC8397EC CRC64;

Query Match 53.8%; Score 42; DB 10; Length 291;  
 Best Local Similarity 66.7%; Pred. No. 45;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MYRASALGSDG 12  
 DB 130 MYROSSLGDSG 141

RESULT 10  
 O9L5F6 PRELIMINARY; PRT; 151 AA.  
 AC O9L5F6;  
 DT 01-DEC-2000 (Tremblrel. 15, Created)  
 DT 01-DEC-2000 (Tremblrel. 15, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE Orf, hypothetical protein.  
 GN R0189.  
 OS Salmonella typhi.  
 OC Plasmid R27.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmonella.  
 CC NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20280091; PubMed=10773089;  
 RA Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,  
 RA Grobeck E., Rose D.J., Taylor D.E.;  
 RT "The complete DNA sequence and analysis of R27, a large, IncHI plasmid  
 from Salmonella typhi that is temperature sensitive for transfer.";  
 RL Nucleic Acids Res. 28:2177-2186(2000).  
 DR EMBL: AF250878; AAF70026.1; -;  
 KW Hypothetical protein; Plasmid.  
 SO SEQUENCE 151 AA; 18075 MW; 15E31FA26518CEB9 CRC64;

Query Match 52.6%; Score 41; DB 2; Length 151;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15  
 DB 107 RYRLNRIGTDGLTV 120

RESULT 11  
 O935M8 PRELIMINARY; PRT; 151 AA.  
 AC O935M8;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)

DE Putative membrane protein.  
 GN HCM1.128.  
 OS Salmonella typhi.  
 OC Plasmid pHCM1.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmonella.  
 RN NCBI\_TaxID=601;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21394947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
 RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrall B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 DR EMBL: AL51383; CAD09726.1; -;  
 KW Plasmid; Complete proteome.  
 SO SEQUENCE 151 AA; 18158 MW; A9927F14B3C3A30F CRC64;

Query Match 52.6%; Score 41; DB 16; Length 151;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15  
 DB 107 RYRLNRIGTDGLTV 120

RESULT 12  
 Q92M00 PRELIMINARY; PRT; 252 AA.  
 AC Q92M00;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)  
 DE Putative oxidoreductase protein (DC 1.1.1.).  
 GN R02563 OR SMC02336.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Rhizobiaceae; Sinorhizobium.  
 CC NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021.  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gonzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,  
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galbert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591791; CAC47142.1; -;  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PROSITE: PS00661; ADH\_SHORT; UNKNOWN\_1.  
 KW Oxidoreductase; Complete proteome.  
 SO SEQUENCE 252 AA; 26560 MW; 0B10B7A0DEFE6628 CRC64;

Query Match 52.6%; Score 41; DB 16; Length 252;  
 Best Local Similarity 57.1%; Pred. No. 57;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYRASALGSDGVRV 15  
 DB 107 RYRLNRIGTDGLTV 120

Db 167 RLHAHAFDGPDIRV 180

RESULT 13  
0930L5 PRELIMINARY: PRT: 258 AA.

AC 0930L5: 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Putative.  
GN RA0180 OR SMA0335.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Plasmid psyma (megaplasmid 1).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OX Rhizobiaceae; Sinorhizobium.  
NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396509; PubMed=11481432;  
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,  
RA Barclay-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,  
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
RT "Nucleotide sequence and predicted functions of the entire  
RT Sinorhizobium meliloti psyma megaplasmid."  
RT Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
DR EMBL: AE007211; AAK64838.1;  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 258 AA; 26146 MW; 446BAHC14A2A8BB1 CRC64;

Query Match 52.6%; Score 41; DB 16; Length 258;  
Best Local Similarity 72.7%; Pred. No. 59;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 ASALGSDGVRV 15  
1 1 1 1 1 1 1 1  
Db 177 ADALGPDGIRV 187

RESULT 14  
081792 PRELIMINARY: PRT: 356 AA.

AC 081792: 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Protein kinase - like protein (Fragment).  
GN F8D20.110.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Koeltter P., Hempel S., Entian K.-D., Hohelsel J., Jesse T.,  
RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;  
RN submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RP [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
CC Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
EMBL: AL031135; CAA20030.1;  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_kinase.  
DR InterPro: IPR004040; STY\_pkinase.

DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00221; STYK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Kinase; Serine/threonine-protein kinase.  
FT NON\_PRT 1  
SQ SEQUENCE 356 AA; 40012 MW; FAD9736384FB8221 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 356;  
Best Local Similarity 53.3%; Pred. No. 84;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYKAGALGAEKRRAT 16  
1 1 1 1 1 1 1 1  
Db 333 RYKAGALGAEKRRAT 347

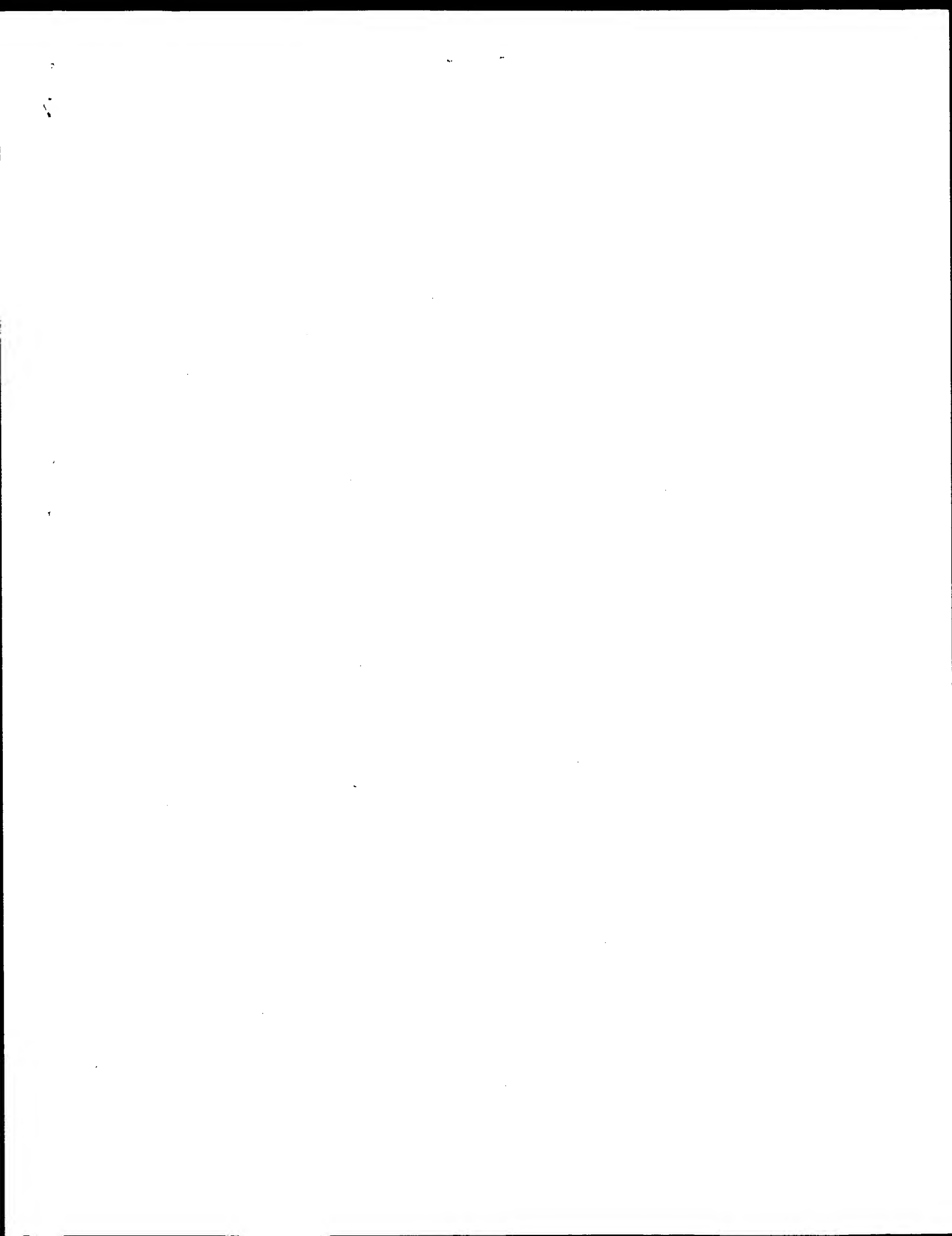
RESULT 15  
09M068 PRELIMINARY: PRT: 415 AA.

AC 09M068: 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Protein kinase-like protein.  
GN AY4G35600.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RN submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RP [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
CC Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
EMBL: AL161587; CAB80276.1;  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_kinase.  
DR InterPro: IPR004040; STY\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00109; TYRKINASE.  
DR PRODOM: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00221; STYK; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 415 AA; 45879 MW; 2FAFD025A219F528 CRC64;

Query Match 52.6%; Score 41; DB 10; Length 415;  
Best Local Similarity 53.3%; Pred. No. 99;  
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 RYKAGALGAEKRRAT 16  
1 1 1 1 1 1 1 1  
Db 392 RYKAGALGAEKRRAT 406

Search completed: January 16, 2003, 10:13:05  
Job time : 31 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:08:27 ; Search time 11 Seconds

(without alignments)  
60.329 Million cell updates/sec

Title: US-09-893-371-2

Sequence: 1 MRRASALGSDGVRRP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSPROT\_40:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	261	1	FABI_ECOLI
2	49	62.8	261	1	FABI_ECOLI
3	45	57.7	260	1	FABI_BUCAT
4	45	56.4	258	1	FABI_ANASP
5	44	56.4	265	1	FABI_PSEAE
6	44	56.4	1391	1	YID5_CAEEL
7	41	52.6	268	1	AMT4_PSEST
8	39	50.0	268	1	FAI12_RHIME
9	38	48.7	109	1	RLAI_MAIZE
10	38	48.7	155	1	CYSC_ARCFU
11	38	48.7	229	1	TPIS_SUITO
12	38	48.7	272	1	FAI1_RHIME
13	38	48.7	424	1	GSA_CAMEL
14	38	48.7	474	1	CYAE_BORPE
15	38	48.7	513	1	YM26_MYCTU
16	37	47.4	545	1	DCIP_AZOBR
17	37	47.4	353	1	HI82_MYCTU
18	37	47.4	551	1	AMT4_PSESA
19	37	47.4	622	1	C2AC_BACTU
20	37	47.4	2212	1	RRPL_EBOZM
21	36	46.2	229	1	HIS4_PYRAE
22	36	46.2	237	1	HPV4_METTF
23	36	46.2	258	1	FABI_SYNY3
24	36	46.2	261	1	FABI_HAENI
25	36	46.2	334	1	BCHT_RHOSH
26	36	46.2	350	1	BCHT_RHOSH
27	36	46.2	361	1	RPL_MYCLE
28	36	46.2	401	1	DXR_STRCO
29	36	46.2	427	1	TOLB_HAENI
30	36	46.2	617	1	ESRI_ICTPU
31	36	46.2	710	1	DAS_PICAN
32	36	46.2	715	1	AT12_HSV1F
33	36	46.2	718	1	AT12_HSV1I

34	36	46.2	778	1	HTR6_HALNI
35	36	46.2	778	1	HTR6_HALSA
36	36	46.2	817	1	PPSA_PYRFU
37	36	46.2	819	1	PPSA_PYRAB
38	36	46.2	821	1	PPSA_PYRHO
39	36	46.2	1004	1	SLPO_BACBR
40	36	46.2	1188	1	PPSA_METJA
41	35	44.9	258	1	FABI_BACSU
42	35	44.9	285	1	GS39_BACSU
43	35	44.9	353	1	WGLM_KCMVK
44	35	44.9	360	1	MTDH_ARATH
45	35	44.9	368	1	RF2_STRCO

## ALIGNMENTS

RESULT 1	FABI_ECOLI	STANDARD:	PRT: 261 AA.
ID	P29132		
AC	01-DEC-1992 (Rel. 24, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-dependent enoyl-ACP reductase).		
DE	FABI OR ENVN OR B1288 OR Z2512 OR ECS1861.		
OS	Escherichia coli, and		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia coli O157:H7.		
NC	NCBI_TaxID=562, 83334;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-30.		
RX	MEDLINE=93123967; PubMed=1364817;		
RA	Bergler H., Hoegenauer G., Turnowsky F.;		
RT	"Sequences of the envn gene and of two mutated alleles in Escherichia coli."		
RL	J. Gen. Microbiol. 138:2093-2100(1992).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=K12 / W3110;		
RA	MEDLINE=94355651; PubMed=8075395;		
RT	Kater M.M., Konigstein G.M., Nijkamp H.J.J., Stultje A.R.;		
RT	"The use of a hybrid genetic system to study the functional relationship between prokaryotic and plant multi-enzyme fatty acid synthetase complexes."		
RL	Plant Mol. Biol. 25:771-790(1994).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	STRAIN=K12 / MG1655;		
RA	MEDLINE=97251357; PubMed=9097039;		
RT	Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horinouchi T.;		
RT	"A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map."		
RL	DNA Res. 3:363-377(1996).		
RN	[5]		
RP	SEQUENCE FROM N.A.		

RC STRAIN=0157:H7 / EDL933 / ATCC 700927;  
RN MEDLINE=210749935; PubMed=11206551;  
RP perna N.T., Plunkett G., Itri, Burland V., Mau B., Glasner J.D.,  
RC MEDLINE=21156231; PubMed=11258796;  
RX Han C.-G., Ohtsuo E., Nakayama K., Kurokawa K., Ishii K., Yokoyama K.,  
RA Hayashi T., Makino K., Onishi M., Murata T., Tanaka M., Tobe T.,  
RA Tida T., Takami H., Honda T., Sasakawa C., Ogatawara N., Yasunaga T.,  
RA Kuhnra S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
RT 0157:H7 and genomic comparison with a laboratory strain K-12".  
RL DNA Res. 8:11-22(2001).  
RN [7]  
RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RX MEDLINE=94164884; PubMed=8119879;  
RA Begler H., Wallner P., Ebeling A., Leitinger B., Fuchsbichler S.,  
RC Aschauer H., Kollenz G., Hoegenauer G., Turnowsky F.;  
RT "Protein Envm is the NADH-dependent enoyl-ACP reductase (FabI) of  
RT Escherichia coli.";  
RL J. Biol. Chem. 269:5493-5496(1994).  
RN [8]  
RP SEQUENCE OF 1-12.  
RC STRAIN-K12 / EMC2;  
RX MEDLINE=97443975; PubMed=9296640;  
RA Link A.J., Robison K., Church G.M.;  
RT "Comparing the predicted and observed properties of proteins encoded  
RT in the genome of Escherichia coli K-12.";  
RL Electrophoresis 18:1259-1313(1997).  
RN [9]  
RP SEQUENCE OF 1-11.  
RC STRAIN-K12 / W3110;  
RX Frutiger S., Hughes G.J., Pasquali C., Hochstrasser D.F.;  
RL Submitted (FEB-1996) to the SWISS-PROT data bank.  
RN [10]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=97113207; PubMed=8953047;  
RA Baldoek C., Kierfety J.B., Seelikhova S.E., Baker P.J., Stuitje A.R.,  
RA Slabas A.R., Hawkes T.R., Rice D.W.;  
RT "A mechanism of drug action revealed by structural studies of enoyl  
RT reductase.";  
RL Science 274:2107-2110(1996).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=99425142; PubMed=104938822;  
RA Ward W.H., Holdgate G.A., Kowseil S., McLean E.G., Paupelt R.A.,  
RA Clayton E., Nichols W.W., Colls J.G., Minshall C.A., Jude D.A.,  
RA Mistry A., Timms D., Camble R., Hales N.J., Britton C.D.,  
RA Taylor I.W.;  
RT "Kinetic and structural characteristics of the inhibition of enoyl  
RT (acyl carrier protein) reductase by triclosan.";  
RL Biochemistry 38:12514-12525(1999).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).  
RX MEDLINE=99329134; PubMed=10398587;  
RA Stewart M.J., Parikh S., Xiao G., Tonge P.J., Kisker C.;  
RT "Structural basis and mechanism of enoyl reductase inhibition by  
RT triclosan.";  
RL J. Mol. Biol. 290:859-865(1999).  
RN [13]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
RX MEDLINE=99215552; PubMed=10201369;  
RA Levy C.W., Roudinskova A., Sedelikhova S., Baker P.J., Stuitje A.R.,  
RA Slabas A.R., Rice D.W., Rafferty J.B.;  
RT "Molecular basis of triclosan activity.";

```

BL      Nature 398:383-384(1999).
CC      -1- CATABOLIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
CC      2,3-dehydroacyl-[acyl-carrier protein] + NADH.
CC      -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
CC      -1- SUBUNIT: HOMOTETRAMER.
CC      -1- SUBCELLULAR LOCATION: Inner membrane-associated.
CC      -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZBORINE INTERFERES WITH THE
CC      ACTIVITY BY BINDING TO THE PROTEIN.
CC      -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
CC      DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; M97219; AAA17755.1; -
CC      EMBL; X78733; CAA55381.1; -
CC      EMBL; AE000227; AAC74370.1; -
CC      EMBL; D90766; BAA14841.1; -
CC      EMBL; D90767; BAA14849.1; -
CC      EMBL; AE005378; AAG56524.1; -
CC      EMBL; AP002556; BAB35284.1; -
CC      PIR; A47681; A47681.
CC      PDB; 1DRG; 28-JAN-98.
CC      PDB; 1DFI; 28-JAN-98.
CC      PDB; 1DFI; 28-JAN-98.
CC      PDB; 1QGG; 21-SEP-99.
CC      PDB; 1QSG; 21-JUL-99.
CC      PDB; 1D8A; 28-OCT-99.
CC      SWISS-2DPAGE; P29132; COLI.
CC      EcoGene; Egl1528; fabI.
CC      InterPro; IPR002198; ADH_short.
CC      Pfam; PF00106; adh_short; 1.
CC      Oxidoreductase; NAD; Fatty acid biosynthesis; Antibiotic resistance;
CC      Inner membrane; 3D-structure; Complete proteome.
CC      KW      inner membrane; 3D-structure; Complete proteome.
CC      FT      INIT_MET 0
CC      FT      NP_BIND 9 35 NAD (BY SIMILARITY).
CC      FT      MUTAGEN 92 92 G->S: DIAZBORINE RESISTANCE
CC      FT      MUTAGEN 240 240 S->F: PRODUCES TEMPERATURE-SENSITIVE
CC      FT      MUTAGEN 240 240 PHENOTYPE.
CC      SQ      SEQUENCE 261 AA; 27732 MW; 434B019F3485956 CRC64;
CC      -----
QY      1 MRYRASALGSDGVRY 15
DB      169 VRYMANMGPEGVRY 183

RESULT 2
FABI_SALTY STANDARD; PRT; 261 AA.
AC PI6657;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NMDH-
DE dependent enoyl-ACP reductase).
GN FABI OR ENVM OR STM1700.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AG701;
RX MEDLINE=90078098; PubMed=2687243;

```

RA Turnowsky F., Fuchs K., Jeschek C., Hoegenauer G.;  
 RT "envm genes of *Salmonella typhimurium* and *Escherichia coli*.";   
 RN J. Bacteriol. 171:6555-6565(1989).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RA MEDLINE=21534948; PubMed=11677609;  
 RX McCelland M., Sanderson K.E., Speleth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated.  
 CC -1- MISCELLANEOUS: THE ANTIBIOTIC DIAZABORINE INTERFERES WITH THE  
 CC ACTIVITY BY BINDING TO THE PROTEIN.  
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M31806; AA27059.1; -;  
 DR EMBL: AE008775; AA120618.1; -;  
 DR PIR: B43729; B43729.  
 DR HSSP: P29132; IDFI.  
 DR StyGene: SG10095; fabI.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR Oxidoreductase: NAD: Fatty acid biosynthesis; Antibiotic resistance;  
 KW Inner membrane; Complete proteome.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT NP\_BIND 9 35 NAD (BY SIMILARITY).  
 FT MUTAGEN 92 92 G->S; DIAZABORINE RESISTANCE.  
 SQ SEQUENCE 261 AA; 27629 MW; 187B91AEB341B773 CRC64;  
 Query Match 62.8%; Score 49; DB 1; Length 261;  
 Best Local Similarity 60.0%; Pred. No. 0.29;  
 Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Oy 1 MRYRASALGSDGVRV 15  
 Db 169 VRYMANAMGPEGVV 183  
 RESULT 3  
 FABI\_BUCAI  
 ID FABI\_BUCAI STANDARD; PRT; 260 AA.  
 AC P57353;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-  
 DE dependent enoyl-ACP reductase).  
 GN FABI OR BU265  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Tokyo 1998;  
 RC MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS.";  
 RL Nature 407:81-86(2000).  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP001118; BAB12975.1; -;  
 DR HSSP: P29132; IDFI.  
 KW Oxidoreductase: NAD: Fatty acid biosynthesis; Complete proteome.  
 FT NP\_BIND 10 36 NAD (BY SIMILARITY).  
 SQ SEQUENCE 260 AA; 28779 MW; DCA04AC740D6DAD CRC64;  
 Query Match 57.7%; Score 45; DB 1; Length 260;  
 Best Local Similarity 53.3%; Pred. No. 1.4;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 MRYRASALGSDGVRV 15  
 Db 170 VRYMANSLGKENVV 184  
 RESULT 4  
 FABI\_ANASP  
 ID FABI\_ANASP STANDARD; PRT; 258 AA.  
 AC Q05069;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-  
 DE dependent enoyl-ACP reductase).  
 GN FABI OR ALA4391.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iritiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium *Anabaena* sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: L10036; AAD04184.1; ALT\_INT.  
DR EMBL: AP003596; BAB76090.1; ALT\_INIT.  
DR HSSP: P29132; IDFI.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short.1.  
KM Oxidoreductase; NAD; Fatty acid biosynthesis; Complete proteome.  
FT NP\_BIND 10 36 NAD (BY SIMILARITY).  
SQ SEQUENCE 258 AA; 27421 MW; 45012CA8497E001B CRC64;

Query Match 56.4%; Score 44; DB 1; Length 258;  
Best Local Similarity 60.0%; Pred. No. 2.1;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALSGDGVY 15  
Db 172 VRYLASELGSQNIYV 186

RESULT 5  
FABI\_PSEAE STANDARD; PRT; 265 AA.  
AC 092FEA; 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Enoyl-[Acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9) (NADH-  
DE dependent enoyl-ACP reductase).  
GN FABI OR PA1806

OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RX MEDLINE=9395061; Pubmed=10464225;  
RA Hoang T.T., Schweizer H.P.;  
RT "Characterization of Pseudomonas aeruginosa enoyl-acyl carrier protein  
RT reductase (FabI): a target for the antimicrobial triclosan and its  
RL role in acylated homoserine lactone synthesis.";  
RL J. Bacteriol. 181:5489-5497 (1999).

[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN-ATCC 15692 / PA01;  
RX MEDLINE=20437337; Pubmed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Seier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964 (2000).

CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-  
CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).  
CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: AF104262; AAC95362.1; -  
DR EMBL: AF004607; AAC05195.1; -  
DR HSSP: P29132; IDFI.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short.1.  
KM Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;  
FT NP\_BIND 10 36 NAD (BY SIMILARITY).  
SQ SEQUENCE 265 AA; 28006 MW; 3E7A94686DF293 CRC64;

Query Match 56.4%; Score 44; DB 1; Length 265;  
Best Local Similarity 53.3%; Pred. No. 2.2;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALSGDGVY 15  
Db 173 VRYLAGSLGAEGRV 187

RESULT 6  
YLD5\_CAEEL STANDARD; PRT; 1391 AA.  
AC 003570; P39218; P34581;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein C38C10.5 in chromosome III.  
GN C38C10.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-Bristol N2.  
RX MEDLINE=94150718; Pubmed=7906398;  
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fullon L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Lattelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38 (1994).

[2]  
RN REVISIONS, AND ALTERNATIVE SPLICING.  
RP Jones S.J.M.;  
RT Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be  
CC produced by alternative splicing.  
CC -1- SIMILARITY: TO YEAST RGT1.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: Z29115; CAA82365.1; -  
DR EMBL: Z19153; CAA82365.1; JOINED.  
DR EMBL: Z19153; CAA79550.1; -  
DR EMBL: Z29115; CAA79550.1; JOINED.  
DR EMBL: Z29115; CAA82366.1; -

DR EMBL: Z19153; CAA82366.1; JOINED.  
 DR EMBL: Z19153; CAA79551.1; JOINED.  
 DR EMBL: Z19153; CAA79551.1; JOINED.  
 DR PIR: S28289; S28289.  
 DR Wormpep: C38C10.5; CE20550.  
 KM Hypothetical protein; Alternative splicing.  
 FT VANSPLIC 671 676 MISSING (IN ISOFORM A).  
 SQ SEQUENCE 1391 AA: 157165 MW: 8702092D4C55A4BF CRC64;  
 Query Match 56.4%; Score 44; DB 1; Length 1391;  
 Best Local Similarity 50.0%; Pred. NO. 13;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MRYRASALGSDGVRV 16  
 Db 1040 IFRASQMGNGCVNAT 1055  
 RESULT 7  
 AMT4\_PSEST STANDARD; PRT; 548 AA.  
 ID AMT4\_PSEST  
 AC P13507;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glucan 1,4-alpha-maltotetrahydrolase precursor (EC 3.2.1.60) (G4-  
 amyase) (Maltotetraose-forming amyase) (Exo-maltotetrahydrolase)  
 DE (Maltotetraose-forming exo-amyase).  
 CN AMP.  
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=316;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=MO-19;  
 RX MEDLINE=89155431; PubMed=2646279;  
 RA Fujita M., ToriGo K., Nakada T., Tsusaki K., Kubota M., Sakai S.,  
 RA Tsujioka Y.;  
 RT "Cloning and nucleotide sequence of the gene (amyP) for  
 maltotetraose-forming amyase from Pseudomonas stutzeri MO-19.";  
 RL J. Bacteriol. 171:1333-1339(1989).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND REVISIONS TO 286-302.  
 RC STRAIN=MO-19;  
 RX MEDLINE=97271999; PubMed=9126844;  
 RA Morishita Y., Hasegawa K., Matsura Y., Katsube Y., Kubota M.,  
 RA Sakai S.;  
 RT "Crystal structure of a maltotetraose-forming exo-amyase from  
 Pseudomonas stutzeri.";  
 RL J. Mol. Biol. 267:661-672(1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF VARIANT GLN-240.  
 RC STRAIN=MO-19;  
 RX MEDLINE=97428332; PubMed=9281429;  
 RA Yoshida Y., Hasegawa K., Matsura Y., Katsube Y., Kubota M.;  
 RT "Crystal structures of a mutant maltotetraose-forming exo-amyase  
 co-crystallized with maltotetraose.";  
 RL J. Mol. Biol. 271:619-628(1997).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF VARIANTS.  
 RC STRAIN=MO-19;  
 RX MEDLINE=20027472; PubMed=10556241;  
 RA Hasegawa K., Kubota M., Matsura Y.;  
 RT "Roles of catalytic residues in alpha-amyases as evidenced by the  
 structures of the product-complexed mutants of a maltotetraose-forming  
 amyase.";  
 RL Protein Eng. 12:819-824(1999).  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-D-glucosidic linkages  
 in amyloseous polysaccharides so as to remove successive  
 maltotetraose residues from the non-reducing chain ends.  
 CC -1- COFACTOR: BINDS TWO CALCIUM IONS.  
 CC -1- PATHWAY: Starch degradation.

CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: THERE ARE SEVERAL ISOZYME FORMS OF THIS PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M24516; AAA25707.1; -.  
 DR PIR: A32803; A32803.  
 DR PDB: 2AMG; 01-APR-97.  
 DR PDB: 1JDA; 15-OCT-97.  
 DR PDB: 1JDC; 15-OCT-97.  
 DR PDB: 1JDD; 15-OCT-97.  
 DR PDB: 1QI3; 24-NOV-99.  
 DR PDB: 1QI4; 24-NOV-99.  
 DR PDB: 1QI5; 24-NOV-99.  
 DR PDB: 1QPK; 17-NOV-99.  
 DR InterPro: IPR000461; Alpha-amyase.  
 DR InterPro: IPR002044; CBD\_4.  
 DR Pfam: PF00128; alpha-amyase; 1.  
 DR Pfam: PF00686; CBM\_20; 1.  
 DR ProDom: PD001568; CBD\_4; 1.  
 KW Hydrolyase; Glycosidase; signal; Carbohydrate metabolism; Calcium;  
 KW 3D-structure.  
 FT SIGNAL 1 21  
 FT CHAIN 22 548 GLUCAN 1,4-ALPHA-MALTOSE HYDROLASE.  
 FT DISULFID 161 171  
 FT DISULFID 227 272  
 FT CA\_BIND 22 22 2.  
 FT CA\_BIND 23 23 2.  
 FT CA\_BIND 34 34 2.  
 FT CA\_BIND 37 37 2.  
 FT CA\_BIND 38 38 2.  
 FT CA\_BIND 137 137 1.  
 FT CA\_BIND 172 172 1.  
 FT CA\_BIND 175 175 1.  
 FT CA\_BIND 183 183 1.  
 FT ACT\_SITE 214 214 1.  
 FT ACT\_SITE 240 240 1.  
 FT ACT\_SITE 315 315 1.  
 FT CONFLICT 286 302  
 FT  
 SQ SEQUENCE 548 AA: 59876 MW: 2887217B379158F CRC64;  
 REF. 1).  
 GSIAWKHGLNGNPDPR -> ARSPTGSPERQSRPA (IN  
 Query Match 52.6%; Score 41; DB 1; Length 548;  
 Best Local Similarity 60.0%; Pred. NO. 16;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MRYRASALGSDGVRV 15  
 Db 116 LRQASALGAGCVKV 130  
 RESULT 8  
 FA12\_RHIME STANDARD; PRT; 268 AA.  
 ID FA12\_RHIME  
 AC P58381;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enoyl-lacyl-carrier-protein) reductase [NADH] 2 (EC 1.3.1.9) (NADH-  
 dependent enoyl-ACP reductase 2).  
 GN FABY1 OR R00246 OR SMC00326  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=1021:  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Drenan S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D.,  
 RA Pohl T., Portetleau D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT *Sinorhizobium meliloti* strain 1021.",  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-  
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.  
 CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).  
 CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN  
 CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL591783; CAC41683.1; -;  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short.1.  
 KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;  
 KW Complete proteome.  
 FT NP\_BIND 11 37 NAD (BY SIMILARITY).  
 SQ SEQUENCE 268 AA; 26671 MW; 83D2127C7D76628C CRC64;  
 Query Match 50.0%; Score 39; DB 1; Length 268;  
 Best Local Similarity 53.3%; Pred. No. 16;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MRYRASALGSDGVRV 15  
 DB 170 VRYLAADYGSRGIRV 184

RESULT 9  
 RLAL\_MAIZE STANDARD; PRT; 109 AA.  
 ID RLAL\_MAIZE STANDARD; PRT; 109 AA.  
 AC P52855; 024414;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 60S acidic ribosomal protein P1 (L12).  
 GN RPLP1.  
 OS Zea mays (Maize).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 CC Panicoideae; Andropogoneae; Zea.  
 CC NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. W22;  
 RA Hamilton D.A., Turcich M.P., Bokhari-Riza A., Mascarenhas J.P.;  
 RN Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 18-21.  
 RP STRAIN=cv. B73; TISSUE=Ear;  
 RX MEDLINE=97422884; PubMed=9276949;  
 RA Bailey-Serres J., Vangala S., Szick K., Lee C.H.;  
 RT "Acidic phosphoprotein complex of the 60S ribosomal subunit of maize  
 RT seedling roots. Components and changes in response to flooding.",  
 RL Plant Physiol. 114:1293-1305(1997).  
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF  
 CC PROTEIN SYNTHESIS (BY SIMILARITY).  
 CC -----

CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL  
 CC SUBUNIT (BY SIMILARITY).  
 CC -1- PTM: PHOSPHORYLATED.  
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U40147; AAA91168.1; -;  
 DR EMBL: U62752; AAB71079.1; -;  
 DR MaizEDB: 84941; -;  
 DR InterPro: IPR001813; 60S\_ribosomal.  
 DR Pfam: PF0428; 60S\_ribosomal.1.  
 KW Ribosomal protein.  
 FT CONFLICT 9 9 R -> T (IN REF. 2).  
 SQ SEQUENCE 109 AA; 11096 MW; 1BFECE8E34391F080 CRC64;  
 Query Match 48.7%; Score 38; DB 1; Length 109;  
 Best Local Similarity 46.7%; Pred. No. 9.1;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 RYRASALGSDGVRV 16  
 DB 9 RYALALLSDGIAIT 23

RESULT 10  
 CYSC\_ARCFU STANDARD; PRT; 155 AA.  
 ID CYSC\_ARCFU STANDARD; PRT; 155 AA.  
 AC 029953;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable adenylisulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-  
 DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-  
 DE phosphotransferase).  
 GN CYSC OR AF0288.  
 OS Archaeoglobus fulgidus.  
 CC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 CC Archaeoglobaceae; Archaeoglobus.  
 CC NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyrleides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Ullrich T.,  
 RA Cotton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,  
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon *Archaeoglobus fulgidus*.",  
 RT Nature 390:364-370(1997).  
 CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.  
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-  
 CC phosphadenylylsulfate.  
 CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. INVOLVED  
 CC IN THE ENERGY-GENERATING SULFUR OXIDATION PATHWAYS OF MANY  
 CC CHEMOAUTOTROPHIC MICROBES.  
 CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.  
 CC -----

```

DR TIGRFAMs: TIGR00419; tm: 1.
DR PROSITE: PS00171; TIM, FALSE_NEG.
KM Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt; Complete proteome.
KW ACT_SITE 93 BY SIMILARITY.
SQ SEQUENCE 229 AA; 24898 MW; 9DCID550A0183525 CRC64;

QY 3 YRASALGSDGVRV 15 48.7%; Score 38; DB 1; Length 229;
|:| ||::|
Db 188 YKAIEFGADGICV 200 Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0

RESULT 12
FAIL_RHIME STANDARD; PRT: 272 AA.
AC P58380;
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Enoyl-[acyl-carrier-protein] reductase [NADH] 1 (EC 1.3.1.9) (NMDH-
GN dependent enoyl-ACP reductase 1).
FABI OR R00898 OR SMC00005.
OC Bacterium meliloti (Sinorhizobium meliloti).
OC Rhizobia; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hudler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Drenth S., Gloux S.,
RA Gohl T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masny D.,
RA Pohl T., Portetalle D., Puehler A., Purrelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NAD(+) = trans-
CC 2,3-dehydroacyl-[acyl-carrier protein] + NADH.
CC -1- PATHWAY: SECOND REDUCTIVE STEP IN FATTY ACID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC -1- SIMILARITY: SOME SIMILARITY TO THE SHORT-CHAIN
CC DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/4sb-sib.ch).
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: AL591785; CAC45470.1; -.
CC InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
KW Oxidoreductase; NAD; Fatty acid biosynthesis; Inner membrane;
KW Complete proteome.
KW NP_BIND 14 40 NAD (BY SIMILARITY).
SQ SEQUENCE 272 AA; 29148 MW; 1E0FA18A22CDBE36 CRC64;

QY Query Match 48.7%; Score 38; DB 1; Length 272;
Best Local Similarity 53.38; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRV 15
|:| |::|
Ob 173 VRYLAVDGDGNGIRV 187

```



RESULT 13  
GSA\_CAMJG STANDARD: PRT: 424 AA.  
ID GSA\_CAMJG  
AC 09P770; 2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) (GSA)  
DE (Glutamate-1-semialdehyde aminotransferase) (GSA-AT).  
CN HEML OR C30853.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCCTC 11168;  
RX MEDLINE=20150912; PubMed=10689204;  
RA Parhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagels K., Karyshev A.V., Moule S., Pallen M.J., Peen C.W.,  
RA Quail M.A., Rajadream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrett B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
CC -1- CATALYTIC ACTIVITY: (S)-4-amino-5-oxopentanoate = 5-  
CC aminolevulinate.  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; second step.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AL139076; CAB73118.1; -.  
CC DR HSSP: P24630; Z6SA.  
CC DR InterPro: IPR000954; AminoTran.3.  
CC DR InterPro: IPR004639; HemL.  
CC DR Pfam: PF00202; AminoTran.3; 1.  
CC DR TIGRPFAMs: TIGR00713; hemL.1.  
CC DR PROSITE: PS00600; AA\_TRANSF\_CLASS.3; FALSE\_NEG.  
CC KMW Porphyrin biosynthesis; isomerase; Pyridoxal phosphate;  
CC Complete proteome. 263 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC FT BINDING 263  
CC SO SEQUENCE 424 AA; 46092 MW; 7811526A5AD57D CRC64;  
Query Match 48.7%; Score 38; DB 1; Length 424;  
Best Local Similarity 72.7%; Pred. No. 39;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

GN CYAE.  
OS Bordetella pertussis.  
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
OC Bordetella.  
OX NCBI\_TaxID=520;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=18323;  
RX MEDLINE=89091151; PubMed=2905265;  
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;  
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-  
RT haemolysin bifunctional protein of Bordetella pertussis.";  
RL HEMO J. 7:3997-4004(1988).  
CC -1- FUNCTION: CYAE IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE  
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).  
CC -1- SUBCELLULAR LOCATION: Outer membrane (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE PRTF FAMILY OF SECRETION PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X14199; CAA32414.1; -.  
CC DR PIR: S02388; BVRCE.  
CC DR InterPro: IPR003423; OEP.  
CC DR Pfam: PF02321; OEP.2.  
CC DR Hemolysis; Transport; Outer membrane; Signal.  
CC KMW Hemolysis; 1 POTENTIAL.  
CC FT SIGNAL 31  
CC FT CHAIN 32 474  
CC SO SEQUENCE 474 AA; 50204 MW; 29A4F21D377FC957 CRC64;  
Query Match 48.7%; Score 38; DB 1; Length 474;  
Best Local Similarity 63.6%; Pred. No. 44;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

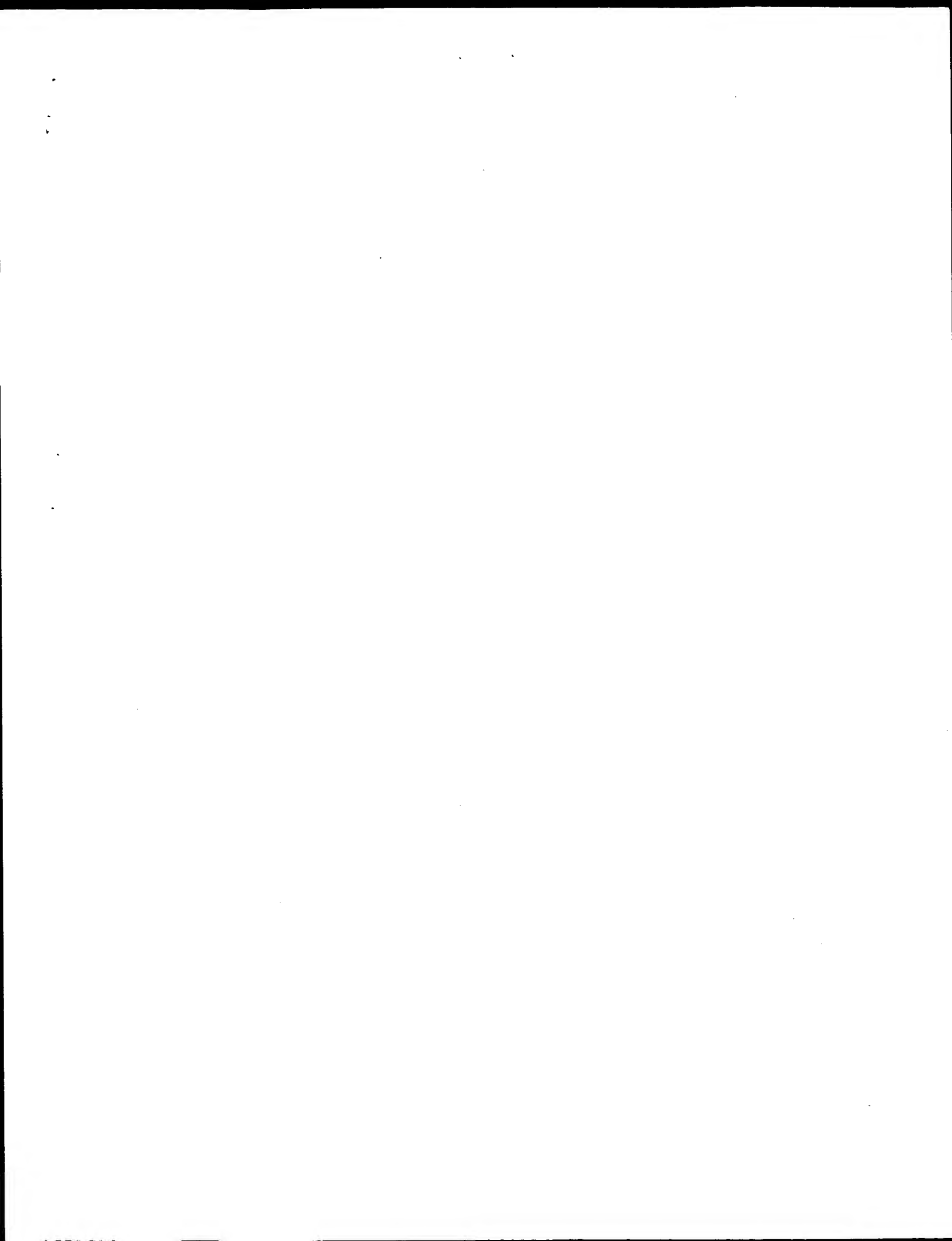
RESULT 15  
YM26\_MYCTU STANDARD: PRT: 513 AA.  
ID YM26\_MYCTU  
AC Q10510;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical protein RV2226.  
DE RV2226 OR MT2285 OR MYCY427.07.  
GN Mycobacterium tuberculosis.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Whitehead S., Barrett B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]



RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khoult H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains."  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: Z70692; CAA94649.1; -;  
 DR EMBL: AE007073; AAK46570.1; -;  
 DR TIGR: MT2285; -;  
 DR TubercuList: RV2226; -;  
 DR Hypothetical protein: Complete proteome.  
 KW CONFLICT 289 D -> N (IN REF. 2).  
 FT SEQUENCE 513 AA: 56333 MW: 8E74BDA8945D7B99 CRC64;

Query Match 48.7%; Score 38; DB 1; Length 513;  
 Best Local Similarity 53.8%; Pred. No. 48;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MRYRASALGSDGV 13  
 :|||:|:|:|:|  
 DB 426 LRYTAATGADNV 438

Search completed: January 16, 2003, 10:12:28  
 Job time : 12 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:11:32 ; Search time 14 Seconds  
(without alignments)  
33.626 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78  
Sequence: 1 MRYRASALGSDGVRYT 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	262	1 US-08-241-766-8	Sequence 8, Appli
2	49	62.8	262	1 US-08-241-766-9	Sequence 9, Appli
3	38	48.7	308	4 US-09-347-803-12	Sequence 12, Appli
4	38	48.7	985	5 PCT-US96-03916-6	Sequence 6, Appli
5	38	48.7	985	5 PCT-US96-03916-6	Sequence 6, Appli
6	37	47.4	623	4 US-09-134-001C-3039	Sequence 3039, Ap
7	37	47.4	623	4 US-09-041-991A-6	Sequence 5, Appli
8	37	47.4	943	4 US-09-134-001C-4009	Sequence 4009, Ap
9	36	46.2	260	4 US-09-134-001C-4009	Sequence 2, Appli
10	36	46.2	537	1 US-08-173-508-2	Sequence 2, Appli
11	36	46.2	537	1 US-08-265-310-2	Sequence 2, Appli
12	36	46.2	537	3 US-08-951-742-2	Sequence 2, Appli
13	36	46.2	635	3 US-09-041-991A-4	Sequence 2, Appli
14	36	46.2	702	6 5240838-15	Sequence 4, Appli
15	36	46.2	3729	2 US-08-804-227C-4	Sequence 4, Appli
16	35	44.9	48	2 US-08-637-759B-480	Sequence 480, App
17	35	44.9	48	3 US-08-871-355A-480	Sequence 480, App
18	35	44.9	48	4 US-09-201-945-480	Sequence 480, App
19	35	44.9	256	4 US-09-292-412-2	Sequence 2, Appli
20	35	44.9	303	4 US-09-420-786A-3	Sequence 3, Appli
21	35	44.9	308	4 US-09-347-803-10	Sequence 10, Appli
22	35	44.9	338	4 US-09-199-637A-47	Sequence 47, Appli
23	35	44.9	350	4 US-08-961-083-126	Sequence 126, App
24	35	44.9	5588	4 US-09-036-987A-6	Sequence 6, Appli
25	35	44.9	5588	4 US-09-370-700-6	Sequence 6, Appli
26	34	43.6	77	4 US-09-465-558-52	Sequence 52, Appli
27	34	43.6	113	1 US-08-211-202-114	Sequence 114, Appli

28	34	43.6	303	2 US-08-286-819A-23	Sequence 23, Appli
29	34	43.6	303	3 US-08-980-357-23	Sequence 23, Appli
30	34	43.6	319	2 US-08-795-927-4	Sequence 4, Appli
31	34	43.6	329	3 US-09-022-669-2	Sequence 2, Appli
32	34	43.6	329	4 US-09-262-749-2	Sequence 2, Appli
33	34	43.6	329	4 US-09-603-567-2	Sequence 2, Appli
34	34	43.6	359	4 US-09-722-139-4	Sequence 4, Appli
35	34	43.6	359	4 US-09-721-832-4	Sequence 4, Appli
36	34	43.6	359	4 US-09-721-832-4	Sequence 4, Appli
37	34	43.6	433	4 US-09-199-637A-349	Sequence 349, App
38	34	43.6	524	3 US-08-529-654-4	Sequence 2, Appli
39	34	43.6	524	4 US-08-957-302A-2	Sequence 2, Appli
40	34	43.6	524	3 US-09-542-403-2	Sequence 2, Appli
41	34	43.6	524	4 US-09-042-709A-20	Sequence 20, Appli
42	34	43.6	623	1 US-08-653-740-7	Sequence 7, Appli
43	34	43.6	623	2 US-09-073-594-7	Sequence 7, Appli
44	34	43.6	623	3 US-09-275-925-7	Sequence 7, Appli
45	34	43.6	633	3 US-09-041-991A-8	Sequence 8, Appli

#### ALIGNMENTS

RESULT 1  
US-08-241-766-8  
Sequence 8, Application US/08241766

Patent No. 5685590

GENERAL INFORMATION:

APPLICANT: JACOBS, W. R.

APPLICANT: COLLINS, D. M.

APPLICANT: BANERJEE, A. M.

APPLICANT: DELISTE, G. W.

APPLICANT: WILSON, T. M.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING

TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/241,766

FILING DATE: 12-MAY-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MONROY, GLADYS H.

REGISTRATION NUMBER: 32,430

REFERENCE/DOCKET NUMBER: 25237-20003.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 262 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-241-766-8

Query Match 62.8% Score 49: DB 1: Length 262:

Best Local Similarity 60.0% Pred. No. 0.51:

Matches 9: Conservative 4: Mismatches 2: Indels 0: Gaps 0:

1 MRYRASALGSDGVRY 15

DB 170 VRYMANMGPBGVRY 184

RESULT 2  
US-08-241-766-9

Sequence 9, Application US/08241766  
Patent No. 5686590

GENERAL INFORMATION:  
APPLICANT: JACOBS, W. R.

APPLICANT: COLLINS, D. M.

APPLICANT: BANERJEE, A. W.

APPLICANT: WILSON, T. M.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING  
TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN LHA AGENT

NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road

CITY: Palo Alto  
STATE: CA

COUNTRY: USA  
ZIP: 94304-1018

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/241,766  
FILING DATE: 12-MAY-1994

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:

NAME: MONROY, GLADYS H.  
REGISTRATION NUMBER: 32,430

REFERENCE/DOCKET NUMBER: 25237-20003.20  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:

LENGTH: 262 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

US-08-241-766-9

Query Match 62.8%; Score 49; DB 1; Length 262;  
Best Local Similarity 60.08; Pred. No. 0.51;

Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 1 MRYASALGSDGVRY 15

DB 170 VRYMANMGPBGVRY 184

RESULT 3  
US-09-347-803-12

Sequence 12, Application US/09347803  
Patent No. 6274379

GENERAL INFORMATION:  
APPLICANT: Famodu, Layo O.

APPLICANT: Hiltz, Bill

APPLICANT: Kinney, Tony

APPLICANT: Orozco, Buddy

TITLE OF INVENTION: Plant Sorbitol Biosynthetic Enzymes  
TITLE REFERENCE: BB-1176

CURRENT APPLICATION NUMBER: US/09/347,803  
CURRENT FILING DATE: 1999-07-02

EARLIER APPLICATION NUMBER: 60/092,952  
EARLIER FILING DATE: July 15, 1998

NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 12  
LENGTH: 308  
TYPE: PRT  
ORGANISM: Oryza sativa  
US-09-347-803-12

Query Match 48.7%; Score 38; DB 4; Length 308;  
Best Local Similarity 48.9%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ASALGSDGV 13

DB 120 ASALGSDGV 128

RESULT 4  
PCT-US96-03916-6

Sequence 6, Application PC/TUS9603916  
GENERAL INFORMATION:

APPLICANT: Wild, Martha A.

APPLICANT: Cochran, Mark D.

TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
TITLE OF INVENTION: AND USES THEREOF

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas

CITY: New York  
STATE: New York

COUNTRY: U.S.A.  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/03916  
FILING DATE: 23-MAR-1995

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/126,597  
FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.

REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 39116-A

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 985 amino acids

TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: Protein  
PCT-US96-03916-6

Query Match 48.7%; Score 38; DB 5; Length 985;  
Best Local Similarity 57.1%; Pred. No. 1.7e+02;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 YRASALGSDGVRY 16

DB 921 YECTVLISDSTRVT 934

RESULT 5  
PCT-US96-03916-66

Sequence 66, Application PC/TUS9603916  
GENERAL INFORMATION:

APPLICANT: Wild, Martha A.

APPLICANT: Wild, Martha A.  
APPLICANT: Cochran, Mark D.  
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/03916  
FILING DATE: 23-MAR-1995  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/126,597  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 39116-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 985 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-03916-66

Query Match 48.7% Score 38; DB 5; Length 985;  
Best Local Similarity 57.1% Pred. No. 1.7e+02;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 YRASALGSDGVRVT 16  
DB 922 YECTVLISDGTTRVT 934

RESULT 6  
US-09-134-001C-3039  
Sequence 3039, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3039  
LENGTH: 309  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3039

Query Match 47.4% Score 37; DB 4; Length 309;  
Best Local Similarity 50.0% Pred. No. 71;  
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 BYRASALGSDGVRV 15  
DB 272 RYEHALGIDPVKL 285

RESULT 7  
US-09-041-991A-6  
Sequence 6, Application US/09041991A  
Patent No. 6107278  
GENERAL INFORMATION:  
APPLICANT: Schnepf, H. Ernest  
APPLICANT: Narva, Kenneth E.  
APPLICANT: Muller-Cohn, Judy  
TITLE OF INVENTION: Toxins Active Against Pests  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,991A  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 51A  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-709  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 623 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-041-991A-6

Query Match 47.4% Score 37; DB 3; Length 623;  
Best Local Similarity 53.8% Pred. No. 1.5e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RASALGSDGVRVT 16  
DB 539 RVSSIGSSITRVT 551

RESULT 8  
US-09-397-885-5  
Sequence 5, Application US/09397885  
Patent No. 6323007  
GENERAL INFORMATION:  
APPLICANT: Moller, Soren  
APPLICANT: Johansen, Charlotte  
APPLICANT: Schaefer, Thomas  
APPLICANT: Ostergaard, Peter Rahbek  
APPLICANT: Hoeck, Lisbeth Hedegaard  
TITLE OF INVENTION: A 2,6,-b-D-fructan Hydrolase Enzyme And  
FILE REFERENCE: 5540.200-US  
CURRENT APPLICATION NUMBER: US/09/397,885  
CURRENT FILING DATE: 1999-09-17  
EARLIER APPLICATION NUMBER: PA 1998 01173

```

; EARLIER FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: PA 1998 01623
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 60/101,615
; EARLIER FILING DATE: 1998-09-24
; EARLIER APPLICATION NUMBER: 60/111,675
; EARLIER FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 943
; TYPE: PRT
; ORGANISM: paenibacillus macerans
; US-09-397-885-5

```

```

Query Match          47.4%; Score 37; DB 4; Length 943;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 MRYRASALGSDG 12
DB 270 LRRRANGCGTDG 281

```

```

RESULT 9
US-09-134-001C-4009
; Sequence 4009, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lyon Doucette-Stamm et al
; TITLE OF INVENTION: EPIDERMIDIS FOR DERMATITIS AND STAPHYLOCOCCUS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4009
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4009

```

```

Query Match          46.2%; Score 36; DB 4; Length 260;
Best Local Similarity 46.7%; Pred. No. 88;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRYRASALGSDGVRV 15
DB 175 VKYLALDGEDNIRV 189

```

```

RESULT 10
US-08-173-508-2
; Sequence 2, Application US/08173508
; Patent No. 5616485
; GENERAL INFORMATION:
; APPLICANT: Bartfield, Daniel
; APPLICANT: Butler, Michael J.
; APPLICANT: Hadary, Dany
; APPLICANT: Jenish, David
; APPLICANT: Krieger, Timothy
; TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
; TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.

```

```

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/173,508
FILING DATE: 23-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/125/CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-173-508-2

```

```

Query Match          46.2%; Score 36; DB 1; Length 537;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRYRASALGSDGVRV 16
DB 6 IRRRAAFGTAGAVT 21

```

```

RESULT 11
US-08-265-310-2
; Sequence 2, Application US/08265310
; Patent No. 5856166
; GENERAL INFORMATION:
; APPLICANT: Bartfield, Daniel
; APPLICANT: Butler, Michael J.
; APPLICANT: Hadary, Dany
; APPLICANT: Jenish, David
; APPLICANT: Krieger, Timothy
; APPLICANT: Malek, Lawrence T.
; APPLICANT: Soostmeyer, Gisela
; APPLICANT: Walczyk, Eva
; APPLICANT: Krysgman, Phyllis
; APPLICANT: Garven, Sheila
; TITLE OF INVENTION: STREPTOMYCES PROTEASES AND IMPROVED
; TITLE OF INVENTION: STREPTOMYCES STRAINS FOR EXPRESSION OF PEPTIDES AND
; TITLE OF INVENTION: POLYPEPTIDES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,310
FILING DATE: 24-JUN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,508
FILING DATE: 23-DEC-1993

```

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 18740/133/CACO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202 672 5300  
TELEFAX: 202 672 5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 537 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-265-310-2

Query Match 46.2%; Score 36; DB 2; Length 537;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRVT 16  
DB 6 IRRRATAFGTAGALVT 21

RESULT 12  
US-08-951-742-2  
Sequence 2, Application US/08951742  
Patent No. 6127144  
GENERAL INFORMATION:  
APPLICANT: Bartfeld, Daniel  
APPLICANT: Michael J. Butler  
APPLICANT: Dany Hadary  
APPLICANT: David Jenish  
APPLICANT: Tim Krieger  
APPLICANT: Lawrence T. Malek  
APPLICANT: Gisela Soostmeyer  
APPLICANT: Eva Walczyk  
APPLICANT: Phyllis Krygsman  
APPLICANT: Sheila Garven  
TITLE OF INVENTION: METHOD FOR EXPRESSION OF PROTEINS IN  
TITLE OF INVENTION: BACTERIAL HOST CELLS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/951,742  
FILING DATE: 16-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 0189740/0140  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 537 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-951-742-2

Query Match 46.2%; Score 36; DB 3; Length 537;  
Best Local Similarity 50.0%; Pred. No. 1.9e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MRYRASALGSDGVRVT 16  
DB 6 IRRRATAFGTAGALVT 21

RESULT 13  
US-09-041-991A-4  
Sequence 4, Application US/09041991A  
Patent No. 6107278  
GENERAL INFORMATION:  
APPLICANT: Schnepf, H. Ernest  
APPLICANT: Narva, Kenneth E.  
APPLICANT: Muller-Cohn, Judy  
TITLE OF INVENTION: Toxins Active Against Pests  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanichik, Lloyd & Saliwanichik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,991A  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-709  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 635 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-041-991A-4

Query Match 46.2%; Score 36; DB 3; Length 635;  
Best Local Similarity 53.8%; Pred. No. 2.3e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 RASALGSDGVRVT 16  
DB 550 RVSSLGNTIVT 562

RESULT 14  
5240838-15  
APPLICANT: LEDEROER, ADRIANUS M.; MAAT, JAN; VERRIPS, CORNELIS  
T.; VISSER, CHRISTIAN; JANOWICZ, ZBIGNIEW A.; HOLLENBERG, CORNELIS P.  
TITLE OF INVENTION: REGULATORY SEQUENCES OF ALCOHOL OXIDASE  
(MOX) AND DIHYDROXYACETONESYNTASE (DAS) OF HANSENUA POLYMORPHA  
NUMBER OF SEQUENCES: 19  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/587,555  
FILING DATE: 24-SEP-1990  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 300,211  
FILING DATE: 23-JAN-1989  
APPLICATION NUMBER: 759,315  
FILING DATE: 26-JUL-1985  
SEQ ID NO: 15  
LENGTH: 702  
5240838-15

Query Match 46.2%; Score 36; DB 6; Length 702;  
Best Local Similarity 77.8%; Pred. No. 2.6e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 SAIGSDGVR 14  
: ||| |||  
Db 277 NALGEDGVR 285

## RESULT 15

US-08-804-227C-4  
Sequence 4, Application US/08804227C  
Patent No. 5876991  
GENERAL INFORMATION:  
APPLICANT: Dehoff, Bradley S.  
APPLICANT: Kuhnstos, Stuart A.  
APPLICANT: Rosteck, Paul R., Jr.  
APPLICANT: Sutton, Kimberly L.  
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THOMAS G. PLANT 1501  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: ASCI(DOS) Text only  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,227C  
FILING DATE: February 21, 1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas, G.  
REGISTRATION NUMBER: 35,784  
REFERENCE/DOCKET NUMBER: X-8231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3729 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-804-227C-4

Query Match 46.2%; Score 36; DB 2; Length 3729;  
Best Local Similarity 66.7%; Pred. No. 1.6e+03;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RYRASALGSDGV 13  
: ||| |||  
Db 1778 RSRAPAGAGDV 1789

Search completed: January 16, 2003, 10:13:48  
Job Time : 15 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:13:13 ; Search time 11 Seconds  
(Without alignments)  
28.914 Million cell updates/sec

Title: US-09-893-371-2  
Perfect score: 78  
Sequence: 1 MRYRASALGSDGVKVT 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB pep: \*  
2: /cgn2\_6/ptodata/2/pubppaa/PCY\_NEW\_PUB pep: \*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_PUB pep: \*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB pep: \*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB pep: \*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB pep: \*  
7: /cgn2\_6/ptodata/2/pubppaa/PCYUS\_PUBCOMB pep: \*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB pep: \*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB pep: \*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB pep: \*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB pep: \*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB pep: \*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB pep: \*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	262	10 US-09-815-242-10152	Sequence 10152, A
2	49	62.8	262	10 US-09-815-242-13824	Sequence 13824, A
3	44	56.4	265	10 US-09-815-242-11851	Sequence 11851, A
4	38	48.7	315	10 US-09-815-242-13289	Sequence 13289, A
5	38	48.7	1233	9 US-09-738-626-4312	Sequence 4312, Ap
6	37	47.4	441	9 US-09-738-626-4275	Sequence 4275, Ap
7	37	47.4	722	9 US-09-738-626-5453	Sequence 5453, Ap
8	37	47.4	1316	9 US-10-120-5444-4	Sequence 4, Appl1
9	37	47.4	1344	9 US-10-120-5444-20	Sequence 20, Appl1
10	37	47.4	1386	9 US-10-120-5444-6	Sequence 6, Appl1
11	36	46.2	50	10 US-09-864-761-40694	Sequence 40694, A
12	36	46.2	295	10 US-09-815-242-11279	Sequence 11279, A
13	36	46.2	358	9 US-09-738-626-4829	Sequence 4829, Ap
14	35	46.2	1160	9 US-10-028-072-234	Sequence 234, App
15	35	44.9	256	10 US-09-815-242-5228	Sequence 5228, Ap
16	35	44.9	256	10 US-09-815-242-12535	Sequence 12535, A
17	35	44.9	256	10 US-09-815-242-12892	Sequence 12892, A
18	35	44.9	316	10 US-09-815-242-13929	Sequence 13929, A

20	35	44.9	350	10 US-09-765-272-126	Sequence 126, App
21	35	44.9	558	10 US-09-770-517C-2	Sequence 2, Appl1
22	34.5	44.2	94	9 US-09-976-059-29	Sequence 29, Appl1
23	34.5	44.2	666	10 US-09-815-242-1933	Sequence 1933, Ap
24	34.5	44.2	677	10 US-09-815-242-10663	Sequence 10663, A
25	34.5	44.2	3034	10 US-09-737-149-25	Sequence 25, Appl1
26	34.5	44.2	3034	10 US-09-737-149-30	Sequence 30, Appl1
27	34.5	44.2	220	10 US-09-815-242-13735	Sequence 13735, A
28	34	43.6	229	9 US-09-965-529-22	Sequence 22, Appl1
29	34	43.6	319	9 US-10-001-426-4	Sequence 4, Appl1
30	34	43.6	319	9 US-10-003-014-4	Sequence 4, Appl1
31	34	43.6	345	10 US-09-925-300-1188	Sequence 1188, Ap
32	34	43.6	423	9 US-09-911-317-4	Sequence 4, Appl1
33	34	43.6	513	10 US-09-833-745-46	Sequence 46, Appl1
34	34	43.6	810	9 US-09-712-363-281	Sequence 281, App
35	34	43.6	911	10 US-09-745-763-140	Sequence 140, App
36	34	43.6	1430	9 US-09-740-274-6	Sequence 6, Appl1
37	34	43.6	3739	9 US-09-860-846-33	Sequence 33, Appl1
38	34	43.6	3739	9 US-09-861-289-33	Sequence 33, Appl1
39	34	43.6	11877	9 US-09-860-846-6	Sequence 6, Appl1
40	34	43.6	11877	10 US-09-861-289-6	Sequence 34943, A
41	33	42.3	87	10 US-09-864-761-34943	Sequence 10428, A
42	33	42.3	220	10 US-09-815-242-10428	Sequence 26, Appl1
43	33	42.3	278	9 US-10-027-806-26	Sequence 26, Appl1
44	33	42.3	278	9 US-10-034-623-26	Sequence 26, Appl1
45	33	42.3	292	9 US-09-738-626-5836	Sequence 5836, Ap

## ALIGNMENTS

RESULT 1  
US-09-815-242-10152  
Sequence 10152, Application US/09815242  
Patent No. US2002061569A1

GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815, 242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191, 078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206, 848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207, 727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242, 578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253, 625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257, 931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269, 308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 10152  
LENGTH: 262  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10152

Query Match 62.8%; Score 49; DB 10; Length 262;  
Best Local Similarity 60.0%; Pred. No. 0.17;







Db 1238 YROSGYGTDCV 1248

## RESULT 12

US-09-864-761-40694

Sequence 40694, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-08-03

PRIOR FILING DATE: 2000-10-04

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVR 15

Db 34 RYRASLASKHISV 47

## RESULT 13

US-09-815-242-11279

Sequence 11279, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

CURRENT FILING DATE: 2001-03-21

Query Match 46.2%; Score 36; DB 10; Length 295;

Best Local Similarity 50.0%; Pred. No. 42;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 RYRASALGSDGVR 15

Db 204 RYMAADLGKGRIV 217

## RESULT 14

US-09-738-626-4829

Sequence 4829, Application US/09738626

Patent No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAOKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738,626  
PRIOR FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4829  
LENGTH: 358  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4829

Query Match 46.2%; Score 36; DB 9; Length 358;  
Best Local Similarity 50.0%; Pred. No. 52;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 YRASALGSDGVRYT 16  
11:1 11:1  
Db 228 YRSSGKGQGVNTT 241

RESULT 15  
US-10-028-072-234  
Sequence 234, Application US/10028072  
Publication No. US2003000431A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Desforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang  
TITLE OF INVENTION:  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24

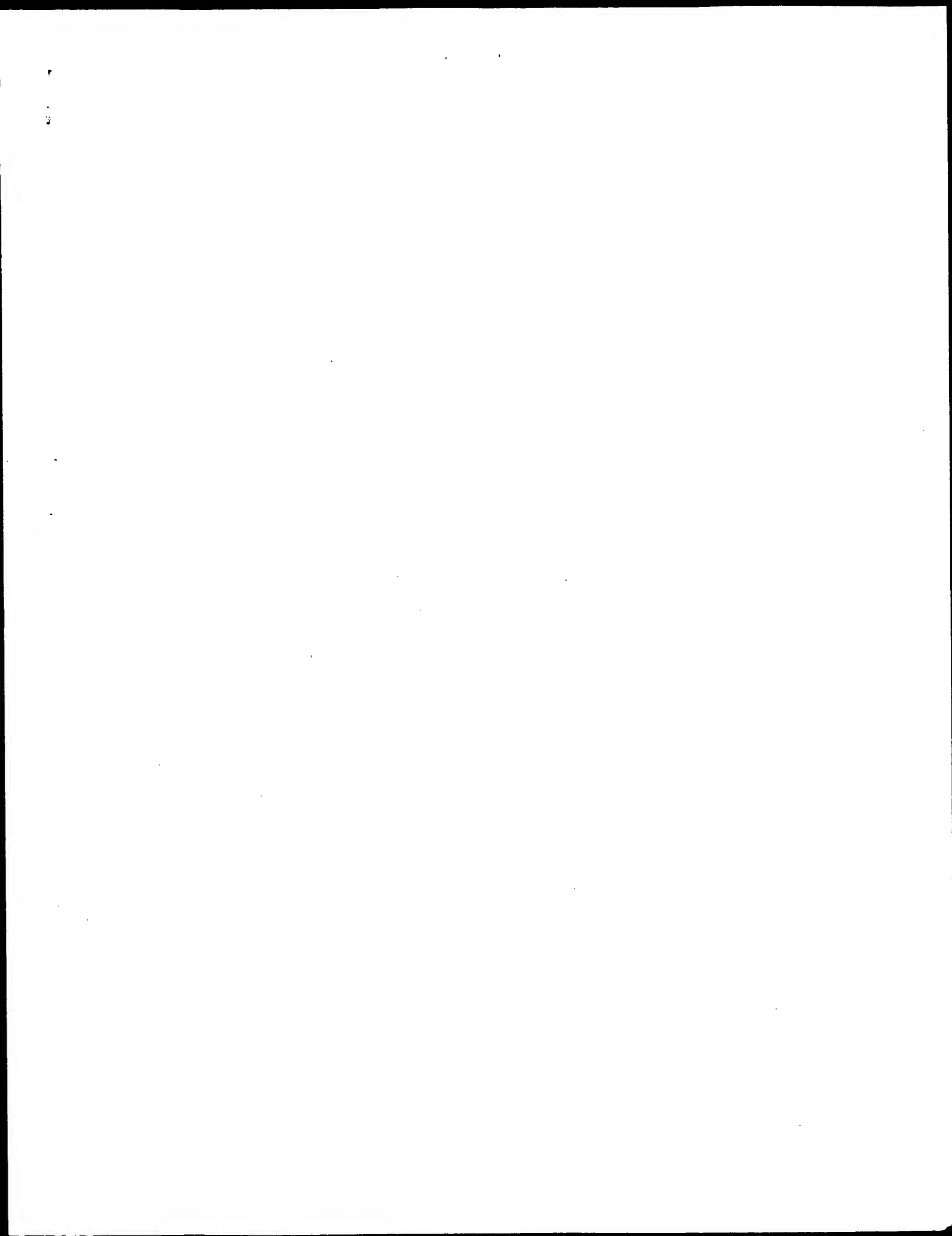
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663

;; PRIOR FILING DATE: 1998-02-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081695  
;; PRIOR FILING DATE: 1998-04-14  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081818  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082999  
;; PRIOR FILING DATE: 1998-04-24  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085149  
;; PRIOR FILING DATE: 1998-05-12  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/086414  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/086430  
;; PRIOR FILING DATE: 1998-05-22  
;; PRIOR APPLICATION NUMBER: 60/087106  
;; PRIOR FILING DATE: 1998-05-28  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088730  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088741  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090538  
;; PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

Query Match 46.2%; Score 36; DB 9; Length 1160;  
Best Local Similarity 77.8%; Pred. No. 2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 RASALGSDG 12  
DB 564 RAGALGADG 572

Search completed: January 16, 2003, 10:16:53  
Job time : 12 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2003, 10:11:12 : Search time 16 Seconds  
(without alignments)  
96.134 Million cell updates/sec

Title: US-09-893-371-2

Perfect score: 78

Sequence: 1 MRYRASALGSDCVKVT 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : PIR\_73:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	62.8	262	1 S48029	enoyl-[acyl-carrie
2	49	62.8	262	1 B43729	enoyl-[acyl-carrie
3	49	62.8	262	2 E90861	enoyl-[acyl-carrie
4	49	62.8	262	2 H85757	enoyl-[acyl-carrie
5	49	62.8	262	2 AD0656	enoyl-[acyl-carrie
6	48	61.5	99	2 S71883	enoyl-[acyl-carrie
7	45	57.7	260	2 G84960	enoyl-[acyl-carrie
8	44	56.4	264	2 AG2354	enoyl-[acyl-carrie
9	44	56.4	265	2 C83419	NADH-dependent eno
10	44	56.4	1112	2 S28289	hypothetical prote
11	44	56.4	1385	2 A88554	protein C38C10.5a
12	44	56.4	1391	2 B88554	protein C38C10.5b
13	42	53.8	291	2 F86451	protein F6N18.2 [i
14	41	52.6	258	2 D95284	probable [imported
15	41	52.6	355	2 T04655	probable serine/th
16	41	52.6	415	2 C85420	protein kinase-lik
17	41	52.6	547	2 A32803	glucan 1,4-alpha-m
18	41	52.6	572	2 T37128	hypothetical prote
19	39	50.0	261	2 C81211	enoyl-[acyl-carrie
20	39	50.0	261	2 E81787	enoyl-[acyl-carrie
21	39	50.0	262	2 AB1196	enoyl-[acyl-carrie
22	39	50.0	262	2 AH1553	enoyl- acyl-carrie
23	39	50.0	274	2 AB2283	hypothetical prote
24	39	50.0	924	2 S75284	chemotaxis protein
25	39	50.0	1296	2 C82521	hemolysin-type cal
26	38.5	49.4	513	2 A90681	probable sensor hi
27	38.5	49.4	513	2 E85531	probable sensor hi
28	38	48.7	109	2 T02716	acidic ribosomal p
29	38	48.7	146	2 T02395	hypothetical prote

30	38	48.7	155	2 H69285	adenylsulfate 3-
31	38	48.7	212	2 E97451	hypothetical prote
32	38	48.7	272	2 AG2669	enoyl-[acyl-carrie
33	38	48.7	312	2 E93164	Hpr(Ser) Kinase/ph
34	38	48.7	316	2 E98030	Hpr(Ser-P) Kinase/p
35	38	48.7	374	2 D86965	probable glycosylt
36	38	48.7	378	2 A70571	probable hexosyltr
37	38	48.7	416	2 D71297	probable flagellar
38	38	48.7	424	2 E81358	glutamate-1-semial
39	38	48.7	437	2 B49933	hnsa export system
40	38	48.7	452	2 AC2718	conserved hypotet
41	38	48.7	474	2 BVBRC2	cyat protein - Bor
42	38	48.7	492	2 G97499	probable periplasm
43	38	48.7	513	2 F70776	hypothetical prote
44	38	48.7	545	1 S44486	indole-3-pyruvate
45	38	48.7	547	2 C84593	probable sugar tra

## ALIGNMENTS

RESULT 1  
S48029  
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) - Escherichia coli (strai  
N:Alternate names: enoyl-ACP reductase; short-chain alcohol dehydrogenase homolog env  
C:Species: Escherichia coli  
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text-change 03-Jun-2002  
C:Accession: S48029; A47681; C64877  
R:Kater, M.M.; Koningsstein, G.M.; Mijkamp, H.J.J.; Stultje, A.R.  
Plant Mol. Biol. 25, 771-790, 1994  
A:Title: The use of a hybrid genetic system to study the functional relationship betw  
A:Reference number: S48029; PMID:94355651; PMID:8075395  
A:Accession: S48029  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <KAT>  
A:Cross-references: EMBL:X78733; NID:9587105; PIDD:CA55381.1; PID:9587106  
A:Experimental source: strain K-12, substrain W3110  
R:Bergler, H.; Hogenauer, G.; Turnowsky, F.  
J. Gen. Microbiol. 138, 2093-2100, 1992  
A:Title: Sequences of the envm gene and of two mutated alleles in Escherichia coli.  
A:Reference number: A47681; PMID:93123967; PMID:1364817  
A:Accession: A47681  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <BPR>  
A:Cross-references: GB:M97219; NID:9145850; PIDD:AA17755.1; PID:9145851  
A:Accession: C64877  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-262 <BLAT>  
A:Cross-references: GB:AE000227; GB:U00096; NID:91787543; PIDD:AC74370.1; PID:917875  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: fabI; envm  
C:Function:  
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the corre  
A:Pathway: fatty acid biosynthesis  
A>Note: Inhibited by palmitoyl-CoA and diazaborine  
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh  
C:Keywords: fatty acid biosynthesis; homotetramer; inner membrane; NAD; oxidoreductas  
F:7-190/Domain: short-chain alcohol dehydrogenase homology <SMDH>  
F:9-35/Region: NAD binding  
Query Match 62.8%; Score 49; DB 1; Length 262;  
Best Local Similarity 60.0%; Pred. No. 0.44;  
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;



C:Accession: 571883  
R:Perito, B.; Allocat, N.; Casalone, E.; Masulli, M.; Dragani, B.; Polistrelli, M.; Acet  
Biochem. J. 318, 157-162, 1996  
A:Title: Molecular cloning and overexpression of a glutathione transferase gene from *P.  
A:Reference number: 571883*; MUID:96358500; PMID:9761466  
A:Accession: 571883  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-99 <P&R>  
A:Cross-references: EMBL:U38482; MUD:q1053074; PIDN:AAC44361.1; PID:q1053075  
C:Experimental source: strain AP 2924  
C:Function:  
A:Description: catalyzes reduction by NADH of enoyl-acyl-carrier-protein to the correspond  
A:Pathway: fatty acid biosynthesis  
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehydr  
C:Keywords: fatty acid biosynthesis; inner membrane; NAD; oxidoreductase

Query Match 61.5%; Score 48; DB 2; Length 99;  
Best Local Similarity 53.3%; Pred. No. 0.24;  
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRYRASALGSDGVRY 15  
Db 7 VRYMANAMGEGIRY 21

RESULT 7  
G84960  
enoyl-[acyl-carrier-protein] reductase (NADH2) (EC 1.3.1.9) [imported] - Buchnera sp. (S  
C:Species: Buchnera sp.  
C:Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text-change 03-Jun-2002  
C:Accession: G84960  
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
A:Reference number: A84930; MUID:20445173; PMID:10993077  
A:Accession: G84960  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-260 <STO>  
A:Cross-references: GB:AP000398; GSPDB:GN00144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: fabI; BU265  
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehydr  
C:Keywords: oxidoreductase

Query Match 57.7%; Score 45; DB 2; Length 260;  
Best Local Similarity 53.3%; Pred. No. 2.3;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MRYRASALGSDGVRY 15  
Db 170 VRYMMSLGKEMIRY 184

RESULT 8  
AG2354  
enoyl-[acyl-carrier-protein] reductase [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence-revision 14-Dec-2001 #text-change 30-Jun-2002  
C:Accession: AG2354  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriuchij  
Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AG2354  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA076090.1; PID:q17133527; GSPDB:GN00179

```

A:Experimental source: strain PCC 7120
C:Genetics:
  A:Gene: all4391
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol dehydrogenase
Query Match          56.4%; Score 44; DB 2; Length 264;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 MRRRASALGSDGVRV 15
      :||| ||| :||
Db 178 VRYLASLGASQNTRV 192

RESULT 9
C83419
NADH-dependent enoyl-ACP reductase PA1806 [Imported] - Pseudomonas aeruginosa (strain ATCC 27803)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83419
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Watterer, P.; Hickey, M.J.; L.
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
  : Loy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: C83419
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <STO>
A:Cross-references: GB:AE004607; GB:AE004091; NID:99947789; PIDN:AMG05195.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
  A:Gene: fabI; PA1806
C:Superfamily: enoyl-[acyl-carrier-protein] reductase (NADH); short-chain alcohol deh
Query Match          56.4%; Score 44; DB 2; Length 265;
Best Local Similarity 53.3%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 MRRRASALGSDGVRV 15
      :||| ||| :||
Db 173 VRYLASLGAEETRV 187

RESULT 10
S28289
hypothetical protein C38C10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jan-2000
R:Thomas, K.
submitted to the EMBL Data Library, December 1992
A:Reference number: S28285
A:Accession: S28289
A:Molecule type: DNA
A:Residues: 1-1112 <THO>
A:Cross-references: EMBL:Z19153
C:Genetics:
  A:Insertions: 412/1; 612/2; 670/3; 729/3; 914/2; 984/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5
Query Match          56.4%; Score 44; DB 2; Length 1112;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 MRRRASALGSDGVRV 16
      :||| :||| :||
Db 1040 IRRFASQMGDGNVAT 1055

RESULT 11
188554

```

protein C38C10.5a [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 17-May-2002  
 C:Accession: A88554  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: A88554  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1385 <STO>  
 A:Cross-references: GB:chr\_III; PIDN:CA79551.1; PID:g3874875; GSPDB:GN00021; CESP:C38C1  
 C:Genetics:  
 A:Gene: C38C10.5a  
 A:Map position: 3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1385;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRVT 16  
 :||| : |||  
 Db 1034 IRRRASQNMNDGVNAT 1049

RESULT 12  
 B88554  
 Protein C38C10.5b [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 02-Aug-2002  
 C:Accession: B88554  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A>Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele  
 A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: B88554  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1391 <STO>  
 A:Cross-references: GB:chr\_III; PIDN:CA82365.1; PID:g3880296; GSPDB:GN00021; CESP:C38C1  
 A>Note: cDNA EST EMBL:CI1839 comes from this gene  
 C:Genetics:  
 A:Gene: C38C10.5b  
 A:Map position: 3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein C38C10.5

Query Match 56.4%; Score 44; DB 2; Length 1391;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 MRYRASALGSDGVRVT 16  
 :||| : |||  
 Db 1040 IRRRASQNMNDGVNAT 1055

RESULT 13  
 FB6451  
 Protein F6N18.2 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: FB6451  
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 Jensen, N.F.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: FB6451  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-291 <STO>  
 A:Cross-references: GB:AE005172; NID:g6714285; PIDN:AAF25981.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F6N18.2  
 A:Map position: 1

Query Match 53.8%; Score 42; DB 2; Length 291;  
 Best Local Similarity 66.7%; Pred. No. 8.9;  
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 MRYRASALGSDG 12  
 :||| : |||  
 Db 130 MRYRSSALGDSG 141

RESULT 14  
 D95284  
 Probable [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymA  
 C:Species: Sinorhizobium meliloti  
 C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: D95284  
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B  
 R; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
 A>Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli  
 A:Reference number: A95262; MUID:2136509; PMID:11481432  
 A:Accession: D95284  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-258 <KUN>  
 A:Cross-references: GB:AE006469; PIDN:AAK64838.1; PID:g14523251; GSPDB:GN00165  
 A:Experimental source: Strain 1021, megaplasmid pSymA  
 R:Galbert, F.; Flanagan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub  
 ler, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeJan  
 hebaull, P.; Vandebol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,  
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: SMA0335  
 A:Genome: plasmid  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 52.6%; Score 41; DB 2; Length 255;  
 Best Local Similarity 72.7%; Pred. No. 12;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 ASALGSDGVRV 15  
 :||| : |||  
 Db 177 ADALGPDGIRV 187

RESULT 15  
 T0465  
 Probable serine/threonine-specific protein kinase (EC 2.7.1.-) F8D20.110 - Arabidopsis  
 N:Alternate names: protein F8D20.110  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-Jan-2000  
 C:Accession: T0465  
 R:Bevan, M.; Rose, M.; Hempel, S.; Ertlan, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mew  
 submitted to the Protein Sequence Database, July 1998  
 A:Reference number: Z15381